Eco RI

1 Kb

2 Kb

NcoI XhoI

3 Kb

4 Kb

XhoI

NcoI

0.65 Kb PCR
FRAGMENT

READING DIRECTION
OF TRANSLATION

NcoI
XhoI

NcoI

5 Kb

Eco RI

FIG._

Nco I　EcoR V　Psp1406 I　Sac I
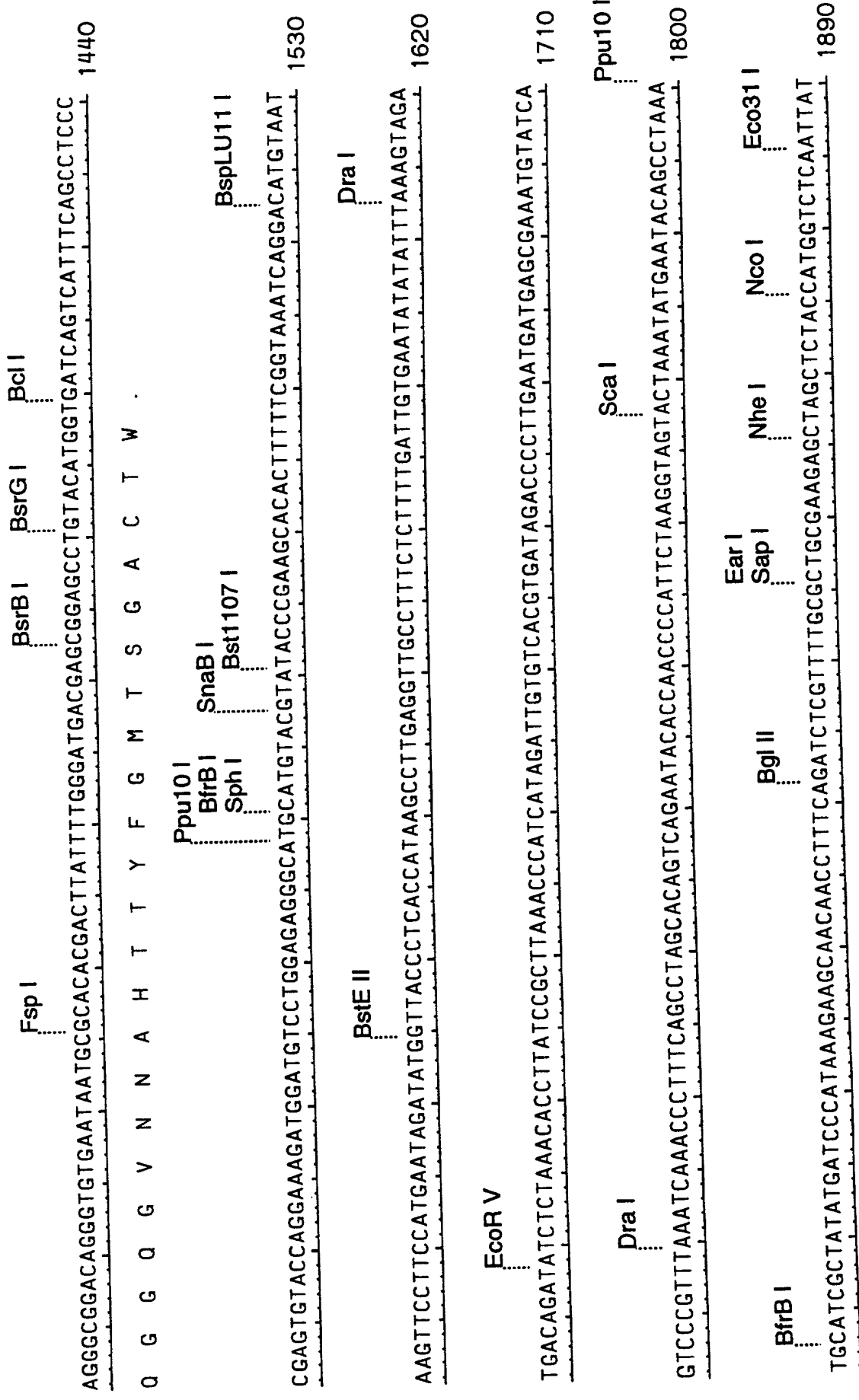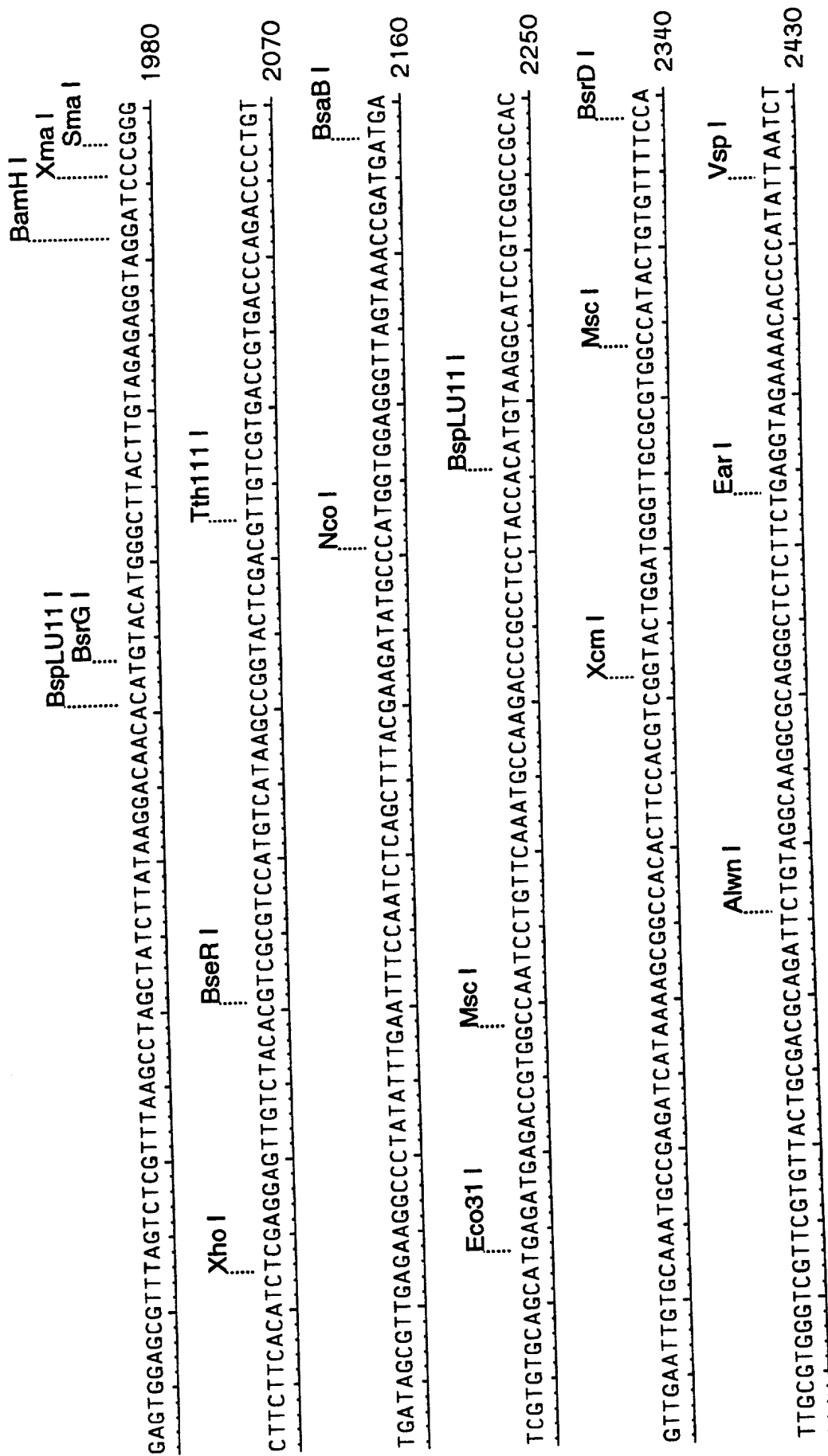
CCATGGTGGTGTCGATATCGGCCAGTAGTCTTTGCCGAAACGTTGAGGGTTACAGTGATCTGCGTCGGACATACTTCGGGGAATCTTACGGC — 90

GGAATATCAAAGTCTTCGGAATATCCATATTGGGAAAGGACAGAAGCTCCGGGGTAGTTTGATAGATGAGCTCCGGTGTATTAAATCGGG — 180

BssH II

AGCTGACAGGAGGTGAGCGTCATGTAGACCATCTAGTAATGTCAGTCGCGCGCCAATTTCGCACATGAAACAAGTTGATTTCGGGACCCCAT — 270

Xho I　Bst1107 I　EclHK I

TGTTACACTCTCTCGGCTACAGCTCGAGATGTGCCTGCCGAGTATACTTAGAAGCCATGCCAGCGTGTTGTTATACGACCAAAGTCAGGG — 360

Pvu I

AATATGAAACGATCGTCGGGATATTTCTTGTGTTTTTATCCTAAATTAGTCTTCCAGTGGTTTATTTAAGAGAGATAGATCCCTTCACAAACACT — 450

Xmn I

CATCCAAACGGACTTCTCATACCACTCATTGACATAAATTTCAAACAGCTCCAGGCGCATTTAGTTCAACATGAAGCAATTCTCCGCCAAAC — 540

signal sequence

M K Q F S A K

FIG. 2A

ACGTCCTCGCAGTTGTGGTGACTGCAGGGCACGCCTTAGCAGCCTCTACGCAAGGCATCTCCGAAGACCTCTACAGCCGTTAGTCGAAA   630

H V L A V V V T A G H A L A A S T Q G I S E D L Y S R L V E

signal sequence

SalI

MscI

TGGCCACTATCTCCCAAGCTGCCTACGCCGACCTGTGCAACATTCCGTCGACTATTATCAAGGGAGAGAAAATTTACAATTCTCAAACTG   720

M A T I S Q A A Y A D L C N I P S T I I K G E K I Y N S Q T

BsaBI

BamHI

ACATTAACGGATGGATCCTCCGGACGACAGCAGCAAAGAAATAATCACCGTCTTCCGTGGCACTGGTAGTGATACGAATCTACAACTCG   810

D I N G W I L R D D S S K E I I T V F R G T G S D T N L Q L

Eco31I

ATACTAACTACACCCCTCACGGCCTTTCGACACCCTACCACCAATGCAACGGTTGTGAAGTACACGGTGGATATTATATTGGATGGGTCTCCG   900

D T N Y T L T P F D T L P Q C N G C E V H G G Y Y I G W V S

FIG. 2B

Tth111 I   BspM I   Acc III

TCCAGGACCAAGTCGAGTCGGCTTGTGTCAAACAGCAGGTTAGCCAGTATATCCGGGACTATGGCGCTGACTGTGACGGGGCCACAGGTATGCCCTCG 990

 V Q D Q V E S L V K Q Q V S Q Y P D Y A L T V T G H S

intron   ApaB I   Pvu II

TGATTTCTTTTCAATTAAGTGTATAAATACTCACTAACTCTACGATAGTCTCGGAGGCGTCCCTGGCAGCACTCACTGCGCCCAGCTGTCTG 1080

                                L G A S L A A L T A A Q L S

intron   BsrG I   Stu I

CGACATACGACAACATCCGCTGTACACCTTCGGCGAACCGCGGCAGCGGCAATCAGGCCTTCGGCGTCGTACATGAACGATGCCTTCCAAG 1170

 A T Y D N I R L Y T F G E P R S G N Q A F A S Y M N D A F Q

Xho I   BspM I   Nco I

CCTCGAGCCCAGATACGACGGCAGTATTTCCGGGTCACTCATGCCAAACGACGGCATCCCAAACCTGCCCGGTGGAGCAGGGGTACGCCC 1260

 A S S P D T T Q Y F R V T H A N D G I P N L P P V E Q G Y A

Sca I

ATGGGCGGTGTAGAGTACTGGAGCGTTGATCCTTACAGCGCCCAGAACACATTTGTCTGCACTGGGGATGAAGTGCAGTGCTGTGAGGCCC 1350

 H G G V E Y W S V D P Y S A Q N T F V C T G D E V Q C C E A

FIG. 2C

FIG. 2D

Fsp I   BsrB I   BsrG I   Bcl I

AGGGCGGACAGGGTGTGAATAATGCGCACACGACTTATTTTGGGATGACGAGGCCTGTACATGGTGATCAGTCATTTCAGCCTCCC   1440

Q G Q G V N N A H T T Y F G M T S G A C T W .

Ppu10 I   SnaB I   BspLU11 I
BfrB I   Bst1107 I
Sph I

CGAGTGTACCAGGAAAGAGATGGATGTCCTGGAGAGGGCATGCATGTACGTATACCCGAAGCACACTTTTCGGTAAATCAGGACATGTAAT   1530

BstE II

AAGTTCCTTCCATGAATAGATATGGTTACCCTCCACCATAAGCCTTGAGGTTGCCTTTCTCTTTGATTGTGAATATATATTTAAAGTAGA   1620

Dra I

EcoR V

TGACAGATATCTCTAAACACCTTATCCGCTTAAACCCCATCATCATAGATTGTGTCACGTGATAGACCCCTTGAATGATGATGAGCGAAATGTATCA   1710

Sca I   Ppu10 I

Dra I

GTCCCGTTTAAATCAAACCCTTTCAGCCTAGCACAGTCAGAATACACCCAACCCCATTCTAAGGTAGTACTAAATATGAATACAGCCTAAA   1800

Ear I   Nhe I   Nco I   Eco31 I
Sap I

BfrB I   Bgl II

TGCATCGCTATATGATCCCCATAAAGAAGCAACAACCTTTCAGATCTCGTTTGCGCTGCGAAGAGCTAGCTCTACCATGGTCTCAATTAT   1890

図 illegible title text

BspLU11 I
BsrG I   BamH I
                                                                          Xma I
                                                                          Sma I

GAGTGGAGGCGTTTAGTCTCGTTTAAGCCTAGCTATCTTATAAGGACAACACATGTACACATGGGCTTACTTGTAGAGAGGTAGGATCCCGGG ─ 1980

Xho I   BseR I   Tth111 I

CTTCTTCACATCTCGAGGAGTTGTCTACACGTCCGTCCATGTCATAAGCCGGTACTCGACGTTGTCGTGACCCAGACCCCTGT ─ 2070

                                                                          Nco I   BsaB I

TGATAGCCGTTGAGAAGGCCCTATATATTTGAATTTCCAATCTCAGCTTTACGAAGAGATATGCCCATGGTGGAGGGTTAGTAAACCGATGATGA ─ 2160

Eco31 I   Msc I   BspLU11 I

TCGTGTGCAGCATGAGATGAGACCGTGGCCAATCCTGTTCAAATGCCAAGACCCGCCTCCTACCACACAGTAAGGCATCCGCTGCGGCCGCAC ─ 2250

          Xcm I   Alwn I   Msc I   BsrD I

GTTGAATTGTGCAAATGCCGAGATCATAAAAAGCGGCCACACTTCCACGTCGGTTGCGCGTGGCCATACTGTGTTTCCA ─ 2340

                                       Ear I   Vsp I

TTGCGTGGGTCGTTCGTTGTTACTGCGACGCGCAGATTCTGTAGGCAAGGGCGCAGGGCTCTCTTCTGAGGTAGAAAACACCCCATATTAATCT ─ 2430

EcoR I
GAATTC ──→ 2436

FIG._2E

CCATGGTGGTGTCGATATCGGCAGTAGTCTTTGCCGAAACGTTGAGGGTTACAGTGATCTGCGTCGGACATACTT
CGGGGAATCTACGGCGGAATATCAAAGTCTTCGGAGCTTGGGAAAGGACAGAAGCTCCGGGTAGTTT
GATAGATGAGCTCCGGGTGTATTAAATCGGGAGCTGACAGGAGCGTGAGCGTCATGTAGACCATCTAGTAATGTCAGT
CGCGCGCAATTCGCACATGCAAGTTGATTTCGGGACCCATTGTTACATCTCTCGGCTACAGCTCGAGATG
TGCCTGCCGAGTATACTTAGAAGCCATGCCAGCCGTGTTGTTATACGACCAAAGTCAGGGAATATGAAACGATCG
TCGGATATTTCTTGTTTTATCCTAAATTAGTCTTCCAGTGGTTTATTTAAGAGATAGATCCCTTCACAAACACT
CATCCAACGGACTTCTCACTCATTGACATAATTTCAAACAGCTCCAGGCGCGCATTTAGTTCAACATGAAAGC
AATTCTCCGCCAAACACGTCCCTGCCAGTTGTGGTGACTGCCACTACTCTCCCAAGCTGCCTACGCCGCAACATTC
CCGAAGACCTCTACAGCCGTTTAGTCGAAAAATTTACAATTCTCAAACTGACATTAACGGATCCTCCCGCGACGACA
CGTCGACTATTATCAAGGGAGAGAAATAATCACCGTCTTCCGTGGCACTGGTAGTGATACGACGGTGATACTACACCC
GCAGCAAAGAAATAATCACCCTACCACGGTGGTCAAATGCAAGCGTGGATAATTCCGGTGATTATTGGATGGGTCTCCG
TCACGCCCTTCGACACCCTACGAGTCGCTTCTTCCGTGTCAAACAGCAGGTTAGCCAGTATCCGGACTATGCGCTGGGGCC
TCCAGGACCAAGTCCCTCGTGATTTCTCTTTCAATTAAGTGTATAAGTGTCGACATACGATAGCTCTTCGGAGCCGTCCC
ACAGGTATGCCCTCGTGCCGCGCCCAGCTGTCTGCGCACATACGACAACATCCGCCTGTACACCTTCGGCGAACCGCGCA
TGGCCAGCACTCACTGCCGCCCCAGCTGTCTGCGCACATGATGCCCTTCCGAGCCCCAGATACGACCGCAGTATTTCC
GCGGCAATCAGGCCCTTCGCGACGGCATCCCAAACCTGCCCCGGTGGAGCAGGGGTACGCCCATGGCGGTGTAGAGT
GGGTCACTCATGCCAACGACGACGCCCTTACAGCGCGCCCCAGAACCACATTTGTCTGCCACTGGGGGATGGAAGTGCAGGGCCC
ACTGGAGCGTTGATCCCTTACAGCGTGAATAATGCGCACCAGGAAAGATGGATGTCCTGGAGAGGGCATGCAGGAGCCGGCATGTACGTGATCAG
AGGGCGGGACAGGGTGTGAATAATGCGCACCAGGAAAGATGTAATAAGTTCCTTCACCGTATACCCATAAGCCTT
TCATTTCAGCCCTCCCGAGTGTACCAGGAACATGTAATAAGTTTTATTTGAATTGTGAATATATTTAAAGTAGATGACAGATCTTCTAAACACCTTATCCGCT
ACACTTTTTCGGTAAATCAGGACACATGTAATAAGTTCCTTCCATTGTTGTGAATATATTTAAAGTAGATGACAGATATCTCTAAACACCTTATCCGCT
GAGGTTGCCTTTCTCTTTTGATTGTGTCACGTGATAGACCCCTTGAATGAGCGAAATGTATCAGTCGTCCCGTTTAAATCA
TAAACCCATCATAGATCGCTAGCACAGTCAGTGCAGAATACACCACCAACCCCATTCTAAGGTAGTACTAAATATGAATACAGCCTAAA
AACCCTTTCAGCTTTACGACAGTCAGTAAAAGAAGCAACAACCTTTCGTTTAGTCTCGTTTGCGCTTCGAGATCGTAGCTCTAC
TGCATCGGCTATATGATCCCCATAAAGAAGCCGTTTAGTTAGTCCTGCTTATAAGGACAACACATGTACATG
CATGGTCTCAATTATGAGTGGAGCGTTAGGAGGATCCCGGCTTCTCTCCGAGGAGTTGTTCTACACGTCGTCCATGTCA
GGCTTACTTGTAGAGAGGTAGGTCGTTGTCGAAGATGTCCCCGGAGGGTTAGTAGCGTTGTGCAGCATGA
TTTCCAATCTCAGCTTTACGACCCGGCCAATCCTGTTCAAATGCCCTCCTACCAGTTCTGATGCGCCCGCCAC
GATGAGACCGGTACGTCGAAATGCCGAGATCATAAAAGCGGCCCACATGTAAGGCGCTTGGATGGGTTGCGTGGCC
GTTGAATTGTGCAAATGCGCAAATCATAAAAGCCGCAGATTCATTCCACCACGTCGGATTTCGTGTTACTGCCTATTTGAA
ATACTGTGTTTTCCATTGCGTGGGTCGTTCGTGTTACTGCCTGGGTGTTAGCTGTGGCC
TCTGAGGTAGAAAAACACCCCATATTAATCTGAATTC

FIG.3

Figure 4

Figure 5

FAE-I3   CCGGCCACGCCCCTCGGCGCCTCCCTGGCGGCACTC   35-mer
FAE-I5   GGCGGCCGAGGGAGTGGCCGGTCACGGTCAGCGGCCGTAGTCC   40-mer

                    i   intron position in original

Y A L T V T G H S   L G A S L A A L
GGACTACGGCGGCTGACCGGTGACCGGCCACTCCCTCGGCGCC   complement, FAE-I5
                CCGGCCACGCCCCTCGGCGCCTCCCTGGCGGCACTC   FAE-I3
Y A L T V T G H A L G A S L A A L

Vector construction

N terminal

actin+ intron
heat shock
senescence

none
FAE excretion
aleurain vac signal
aleurain apo signal
sialyltransf golgi signal

Gene

active glyc site
ser to ala
32aa clip site
codon optimisation

C terminal

none
linker + stop codon
linker +KDEL
linker + frameshift

Promoter | N- signal

CaMV35S | hpt | nos | FAE gene | C signal | Amp

Fig 6

Figure 2

SspI (4063)
XmnI (3858)
SspI (443)
SacI (658)
KspI (665) SacII (665)
NotI (671) EagI (671) XmaIII (671)
XbaI (678)
SpeI (684)
BamHI (690)
SmaI (697)
PstI (705)
EcoRI (707)
EcoRV (715)
BglII (843)
BamHI (875)

AMP

M13 Forward
T7

pCOR105
4173 bp

intron1

Transcription start

Rice Act1 Prom

M13 Reverse   T3

SacI (1077)

EcoRI (1577)

NdeI (1942)   NdeI (1624)

XhoI (1953)

KpnI (1972)

Figure _8_

Generic
ALE-TER
vector

1676 bp
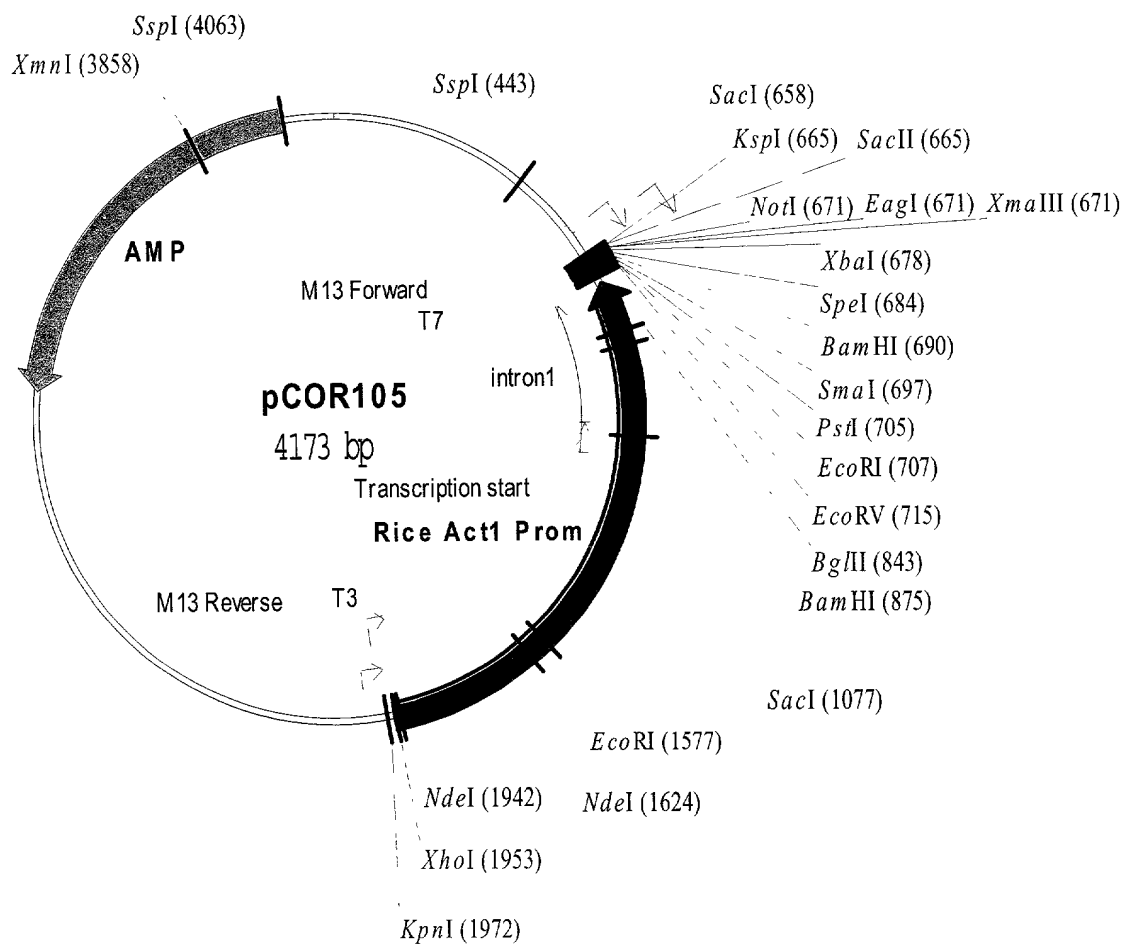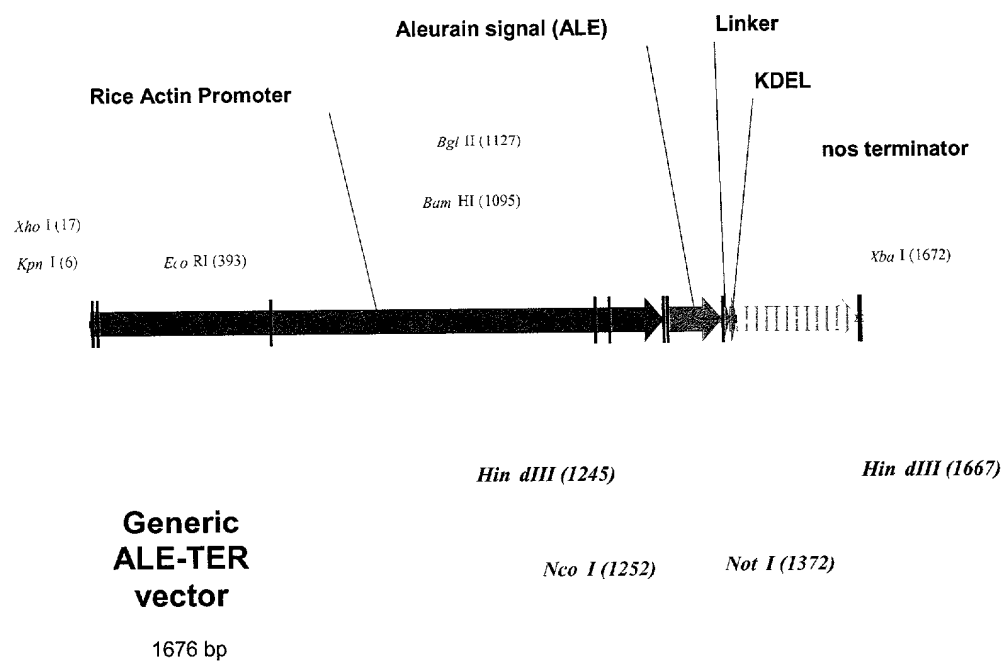
Figure 9

KDEL-COOH ER retention sequence

```
       NotI
       ~~~~~~~~
       A   A   A   K   P   L   K   D   E   L   *
1   GCGGCCGCGA AACCACTGAA GGATGAGCTG TAA
```

Figure 10

FAE-LINKER-FRAMESHIFT structure and sequence


```
        SphI(9)            NotI(24)
          FAE           LINKER   ORF              FRAMESHIFT
```

FAE-LINKER-FRAMESHIFT
48 bp

```
+1   G   A   C   T   W   P   V   A   A   A   E   T   T   E   G
         SphI                        NotI
         ~~~~~~                      ~~~~~~~~

1   GGCGCATGCA  CCTGGCCGGT  CGCGGCCGCG  GAAACCACTG  AAGGATGA
    CCGCGTACGT  GGACCGGCCA  GCGCCGGCGC  CTTTGGTGAC  TTCCTACT
```
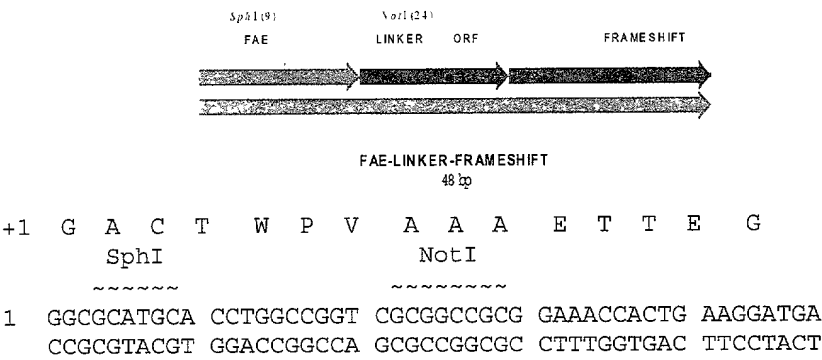
Plant transformation cassettes

Initial vectors			Target	Actin (+ hyg)	H.S.	See1
Original Actin	HS	(diagram)				
		original Actin + hyg			UB 8.1	
TP11.1	TT3	TR 9.4	VAC	UH4	UK3	-
TT5	-	TT5.5	APO	UH6	UH12	-
UA4.4	-	-	APO	UH7	UH13	-
TP8.5	-	-	VAC	UH5	UK 6	-
TP3.1	-	TR8 (-glycos)	VAC	HOX3	UC5.1	-
TU4	-	-	VAC/ER	UH3	UK2	-
TU5	-	-	E.R.	UH8	UH10	-
UG	-	-	E.R.	UH9	UH11	-
TP5.1	TT2	TR6.1	E.R.	UF1	UK1	-
TP4	-	TR2	APO	-	-	-
TP3.1			GOLGI	pJQ4.9 *	-	-
TP3.1			APO	pJQ3.2 *	-	-
TP3.1			VAC	pJ06.3 *	-	pJQ5.2

Legend:
- Aleurain NPIR
- Aleurain NPGR
- RST signal
- PPI signal
- Aspergillus signal
- FAE
- Stop codon
- Linker frameshift
- Linker KPLKDEL

* - Modified actin promoter (Kpn1-EcoR1 deletion and restored NCO site)

Figure 11

Vectors

Original Actin promoter in pCOR105

	Target	Signal sequences	Vectors
(i)	APO	- aleurain-NPGR-FAE	pUH6, pTT5, TT5.5, pTT5.1
		- aleurain-delNPIR -FAE	pUH7, pUA4.4,
(ii)	ER	- aleurain-NPGR-FAE-linker-KDEL	pTU5, pUH8,
		- aleurain-delNPIR-FAE-linker-KDEL	pUG4, pUH9,
(iii)	VAC	- aleurain-NPIR-FAE	pTP11.1, pTR9.4, pUH4, pUK3,
(iv)	ER/VAC	- aleurain-NPIR-FAE-linker-KDEL	pTU4, pUH3,
(v)	VAC	- aleurain-NPIR-FAE-linker-frameshift	pUA1K3, pTP3.1, pUC5.11
(vi)	VAC	- aleurain-NPIR-FAE-linker-stop	pTP8.5, pUH5
(vii)	ER	- Aspergillus signal -FAE-KDEL	pTP5.1, pTP6.1, pUF1,

Modified actin promoter (Kpn1-EcoR1 deletion and restored NCO site)

	Target	Signal sequences	Vectors
(i)	VAC	- aleurain-NPIR-FAE-linker-frameshift	pJ06.3
(ii)	GOLGI	- RST-FAE-linker-frameshift	pJQ3.2
(iii)	APO	- PPI-FAE-linker-frameshift	pJQ4.9

Heat-shock promoter

	Target	Signal sequences	Vectors
(i)	APO	- aleurain-NPGR-FAE	pUH12
		- aleurain-delNPIR-FAE	pUH13
		- Aspergillus signal-FAE	pTP4a2, pTR2.22,
(ii)	ER	- aleurain-NPGR-FAE-linker-KDEL	pUH10
		- aleurain-delNPIR-FAE-linker-KDEL	pUH11
(iii)	VAC	- aleurain-NPIR -FAE	pUK3,pTT3
(iv)	ER/VAC	- aleurain-NPIR-FAE-linker-KDEL	pUK2
(v)	VAC	- aleurain-NPIR-FAE-linker-frameshift	pUC5.11, pHOX3
(vi)	VAC	- aleurain-NPIR-FAE-linker-stop	pUK6
(vii)	ER	- Aspergillus signal -FAE-KDEL	pUK1, pTT2

Senescence promoter

	Target	Signal sequences	Vectors
(i)	APO	- See1-PPI-FAE-linker-frameshift	pJQ5.2
(ii)	VAC	- See1-aleurain-deleted NPIR-FAE	pUB8.1

Figure 12

Figure 13

ALEURAIN -NPIR (Vacuolar) and NPGR (Apoplast) structure and sequence

NPIR UNDERLINE
NPGR BOLD
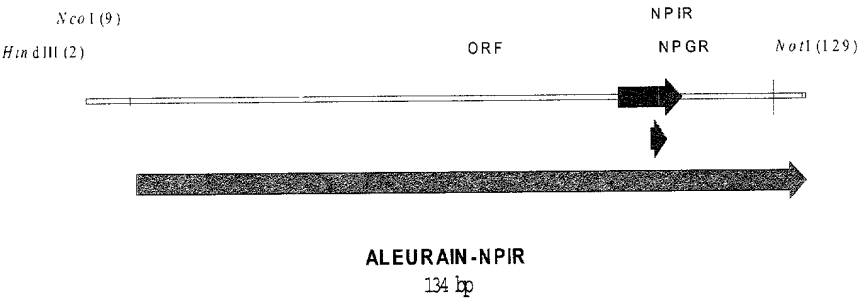
ALEURAIN-NPIR
134 bp

```
+1            M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A
      HindIII NcoI
      ~~~~~~ ~~~~~~~
1  AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG
   TTCGAATGGT ACCGGGTGCG GGCGCAGGAG GAGGACCGCG AGCGGCACGA CCGGTGCCGG CGGCAGCGGC


+1 V   A   S   S   S   S   F   A   D   S   N   P   I   R   P   V   T   D   R   A   A
                                                                           NotI
                                                                           ~~~~~~~~
71  TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGCGG CCGC
    AGCGGAGGAG GAGGAGGAAG CGGCTGAGGT TGGGCTAGGC CGGGCAGTGG CTGGCGCGCC GGCG
```

Figure 14

RAT SIALYL TRANSFERASE **Golgi signal sequence**

```
       HindIII
       ~~~~~~
                    M    I   H   T    N   L   K    K   K   F   S    L   F   I    L   V   F    L   L   F   A
   1   AAGCTTACCA TGATCCACAC CAACCTCAAA AAGAAGTTCT CCCTCTTCAT CCTCGTCTTC CTCCTCTTCG


       ·   V   I   C    V   W   K    K   G   S   D    Y   E   A    L   T   L    Q   A   K   E    F   Q   M  ·
  71   CCGTGATCTG CGTGTGGAAG AAGGGCTCCG ACTACGAGGC CCTCACCCTC CAAGCCAAGG AGTTCCAAAT


       NotI
       ~~~~~~~~
       ·  A   A
 141   GGCGGCCGC
```

Figure 15

POTATO PROTEASE INHIBITOR II Apoplast signal sequence

```
    HindIII
    ~~~~~~
                M   X   V   H   K   E   V   N   F   V   A   Y   L   L   I   V   L   G   L   L   L
  1  AAGCTTACMA TGGMCGTGCA CAAGGAGGTS AACTTCGTSG CCTACCTCCT GATCGTSCTC
     GGCCTCCTCT

                      NcoI
                      ~~~~~~
            ·   L   V   S   A   M   E   H   V   D   A   K   A   C   T   X   E   C   G   N   L
            G   F   G   ·
 71  TGCTCGTSTC CGCCATGGAG CACGTGGACG CCAAGGCCTG CACCCKCGAG TGCGGCAACC
     TCGGCTTCGG

                      NotI
                      ~~~~~~~~
            ·   I   C   P   A   A   A
141  CATCTGCCCG GCGGCCGCC
```

Targeting expression of gfp to different cell compartments

Actin promoter targeting vectors

Legend:
- Actin promoter
- Aleurain signal (NPIR)
- Aleurain signal (NPGR)
- linker + KPLKDEL
- silyl transferase
- Aspergillus signal
- Aleurain signal I (del - NPIR)
- linker + stop codon
- Potato protease inhibitor
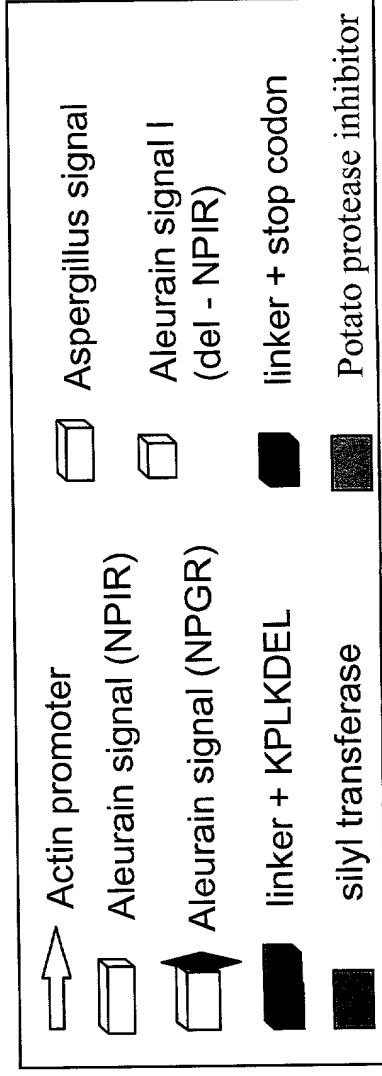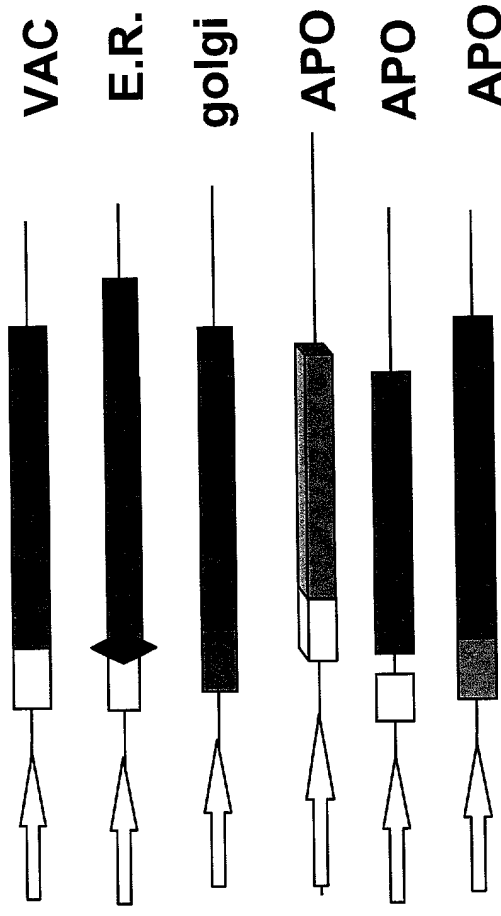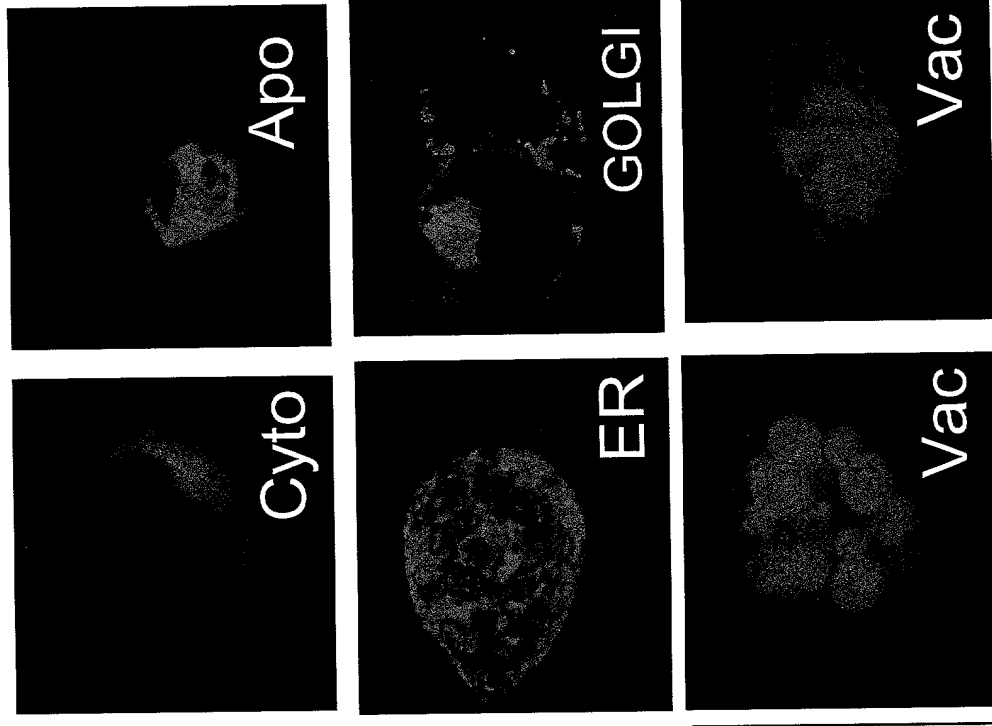
Figure 16

FAE activity in transgenic *Festuca arundinacea* leaves of different ages under ER and APO targeting sequence.
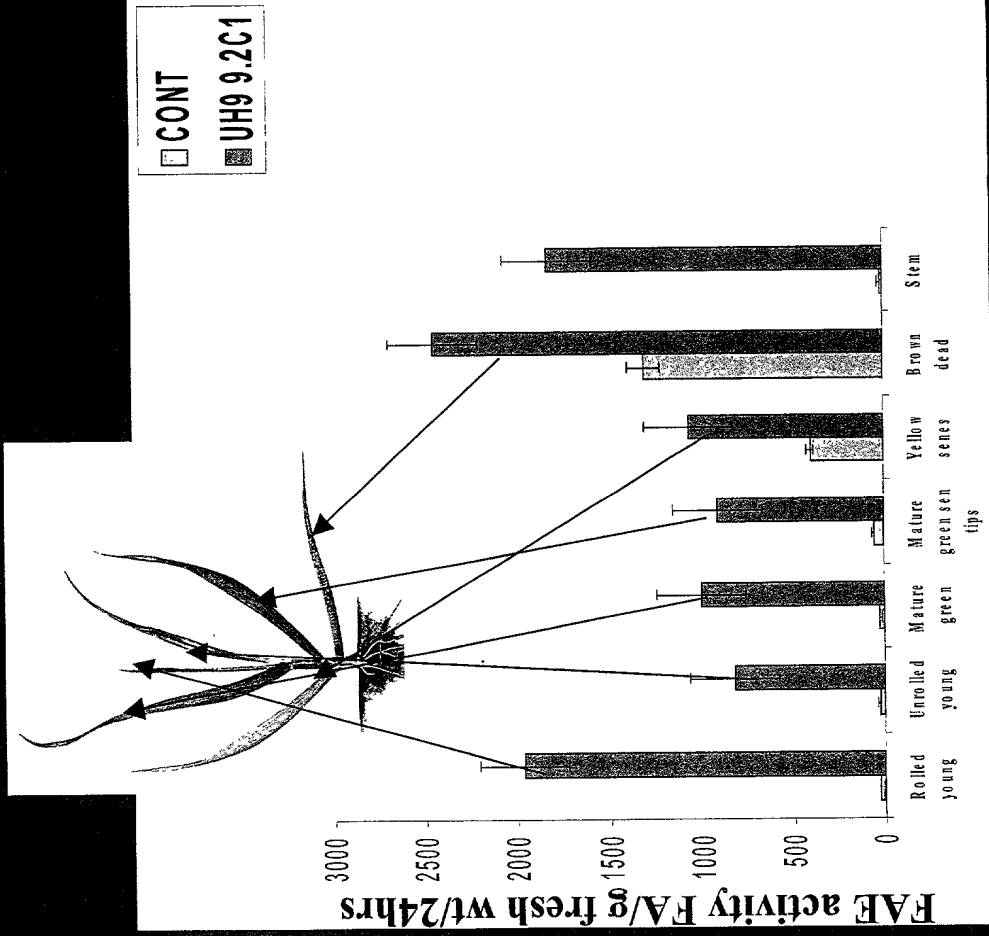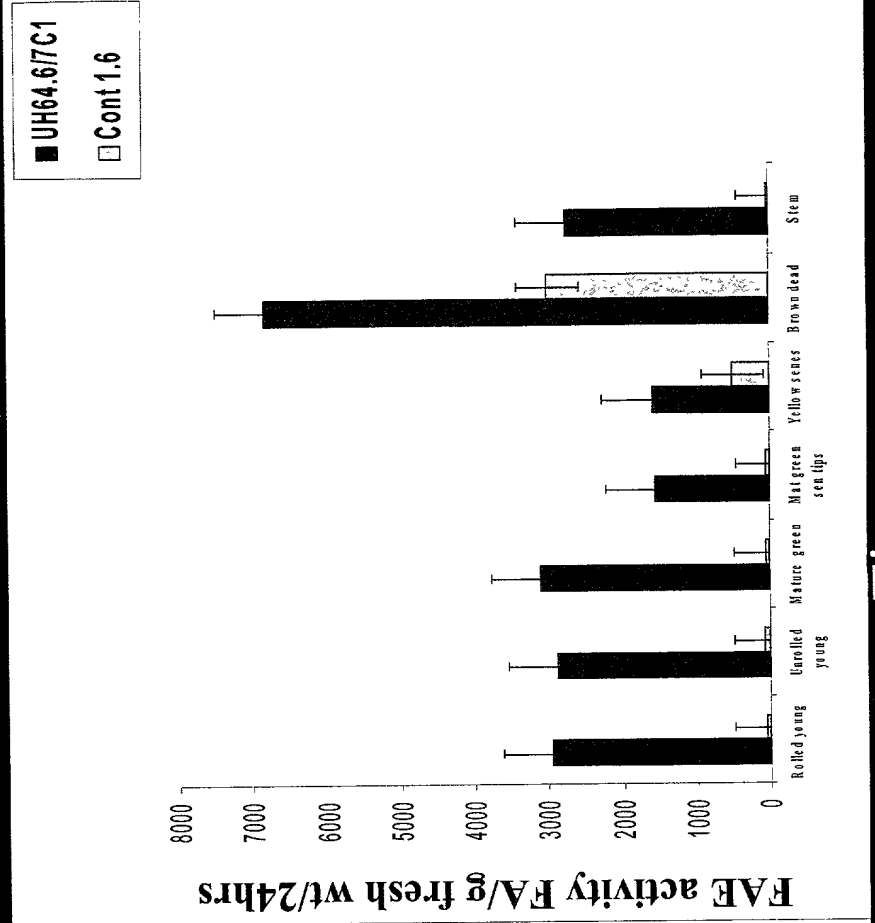
Figure 17

FAE activity in leaves of primary transformants of *Festuca arundinacea* under Vac targeting sequence
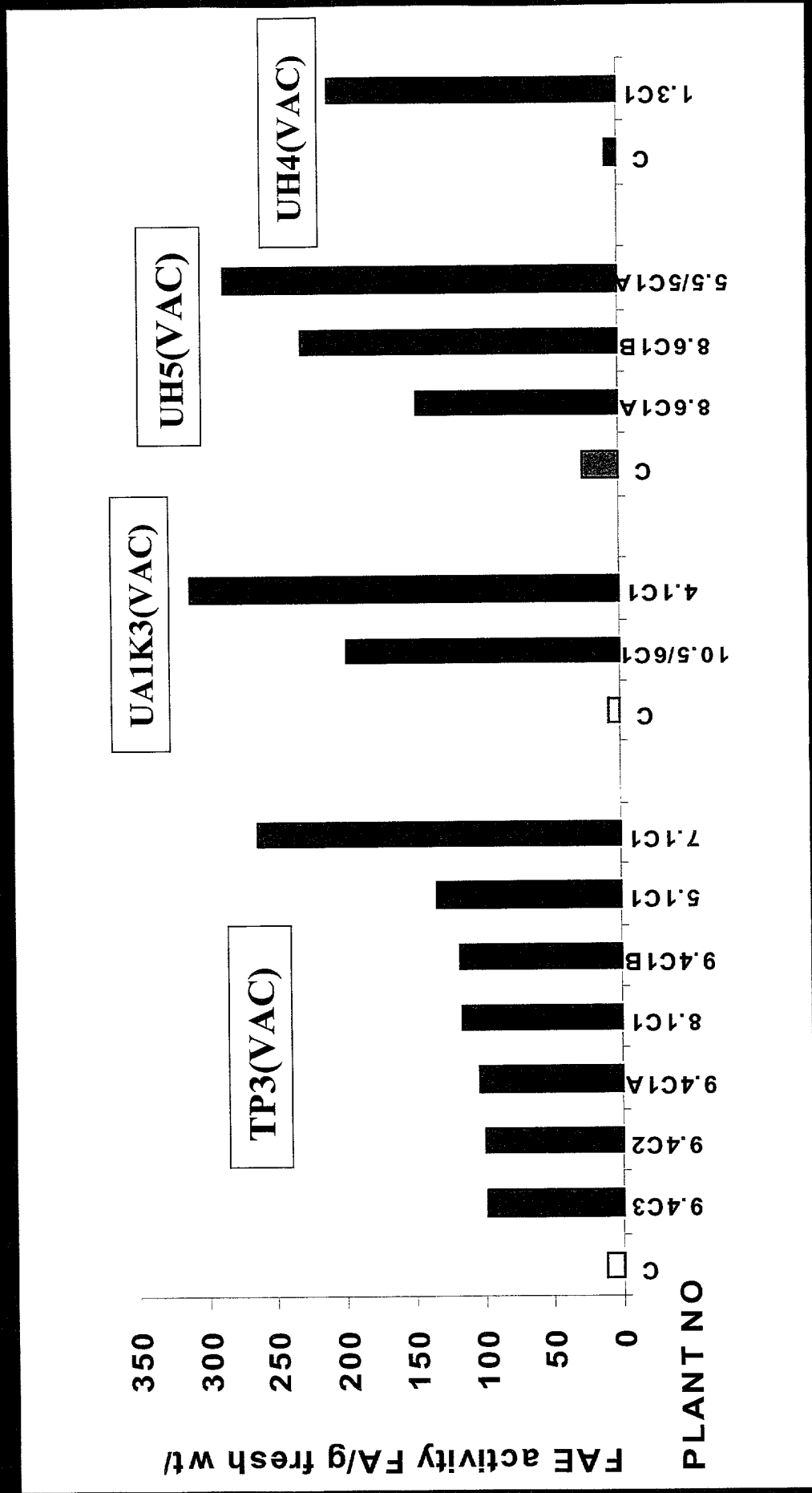
Figure 18

FAE activity in *Lolium mutiflorum* leaves of different ages

Figure 19

FAE activity in leaves of primary transformants of *Lolium multiflorum* under VAC APO and ER targeting sequence.
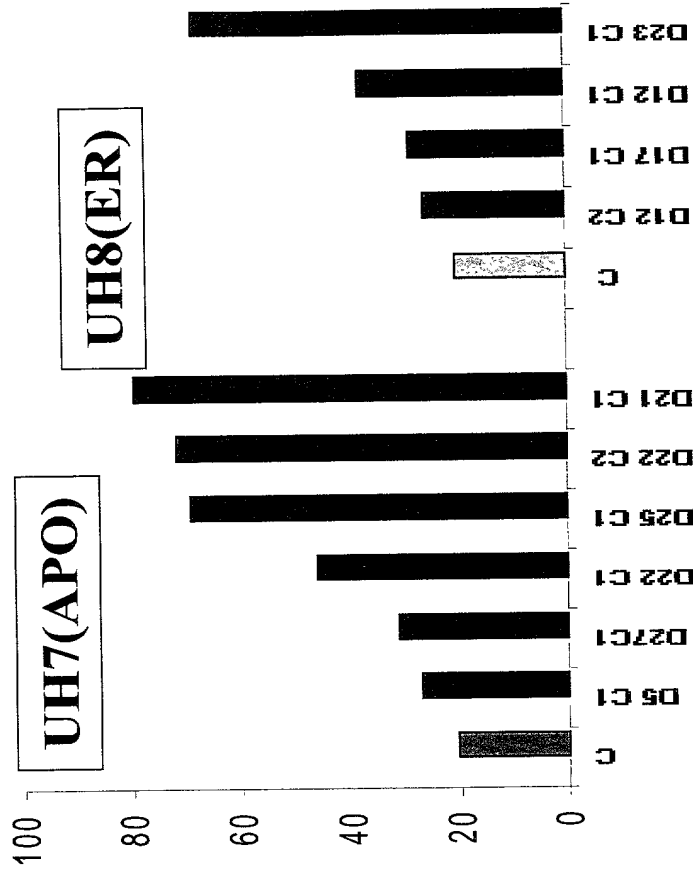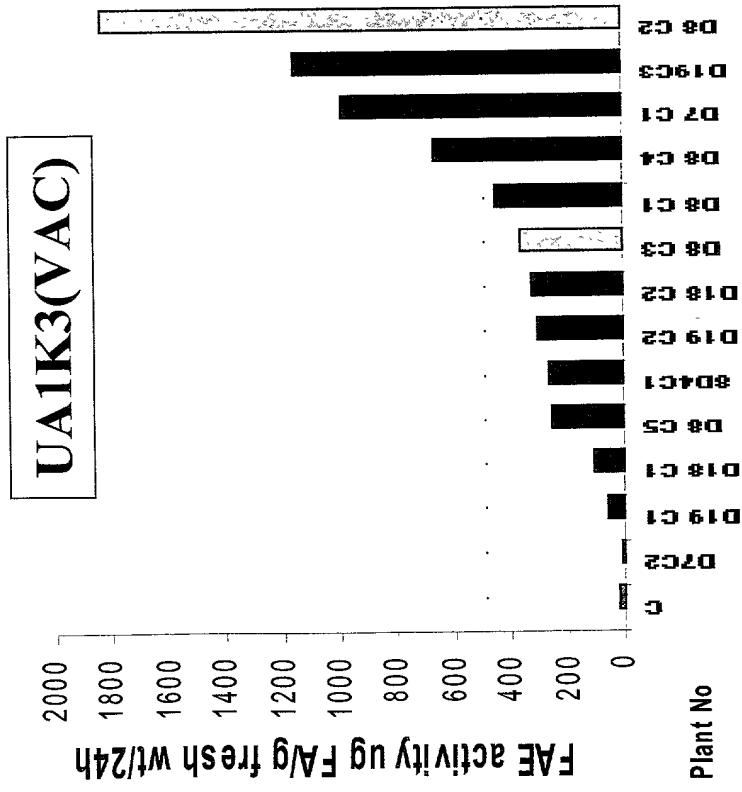
Figure 20

Levels of esterified monomeric and dimeric hydroxycinnamicacids in *Festuca arundinacea* plants expressing FAE under VAC targeting sequence.
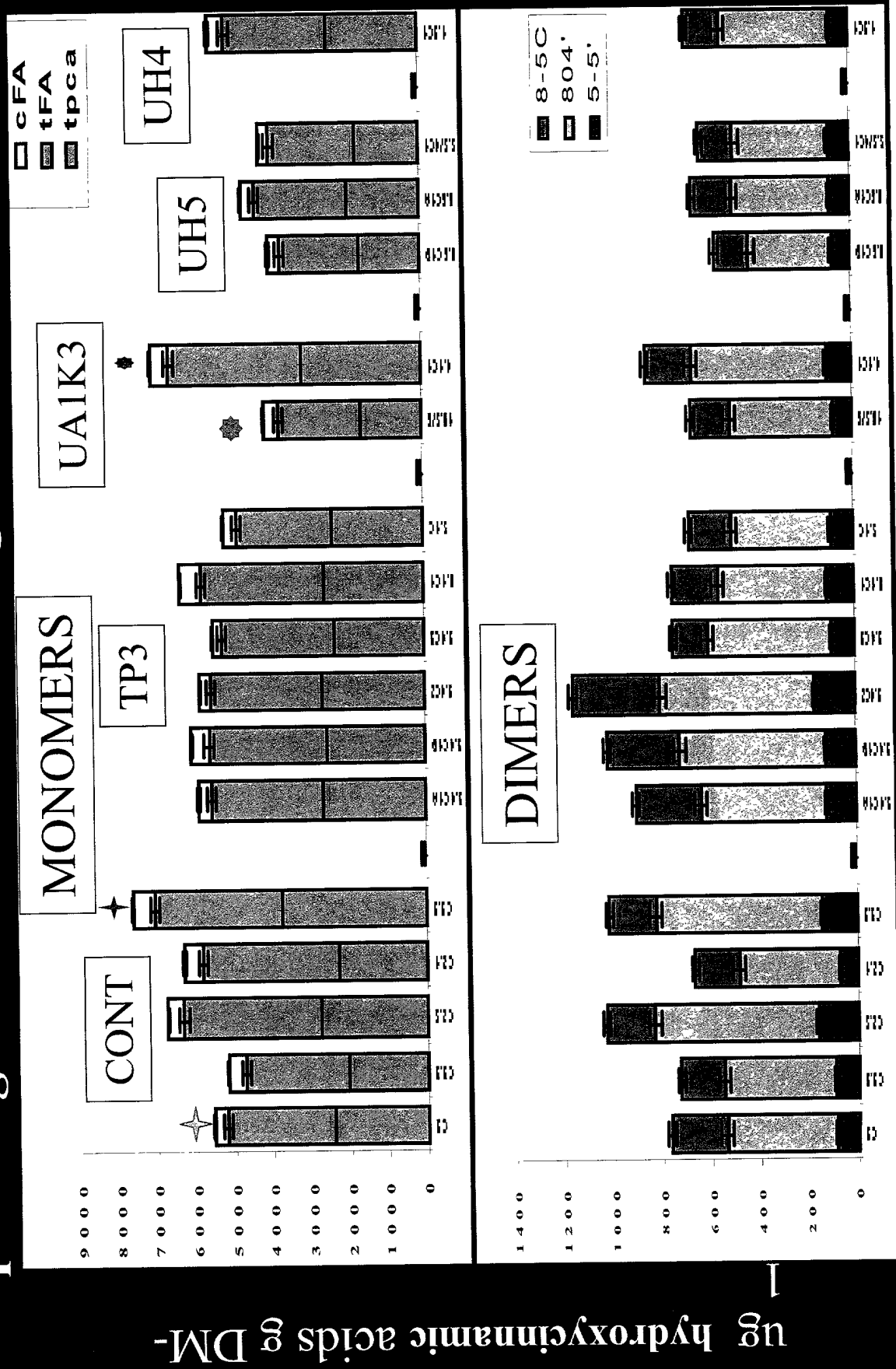
Figure 21

Levels of esterified monomeric and dimeric hydroxycinnamic acids in leaves of *F.a.* expressing FAE under ER and APO targeting
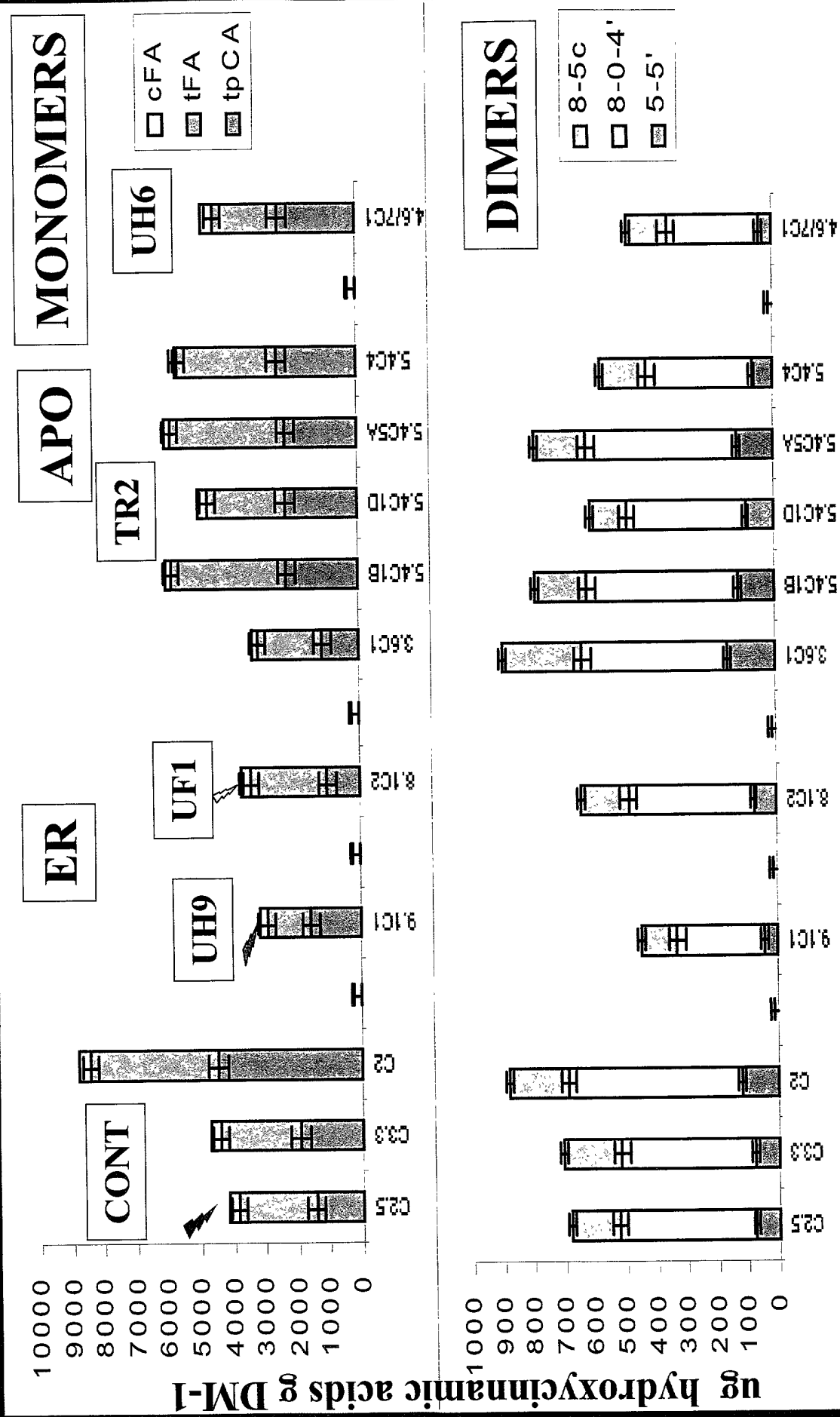
Figure 27

In vitro dry mater digestibility of leaf tissue of mature *Festuca arundinacea* plants expressing FAE under an actin promoter
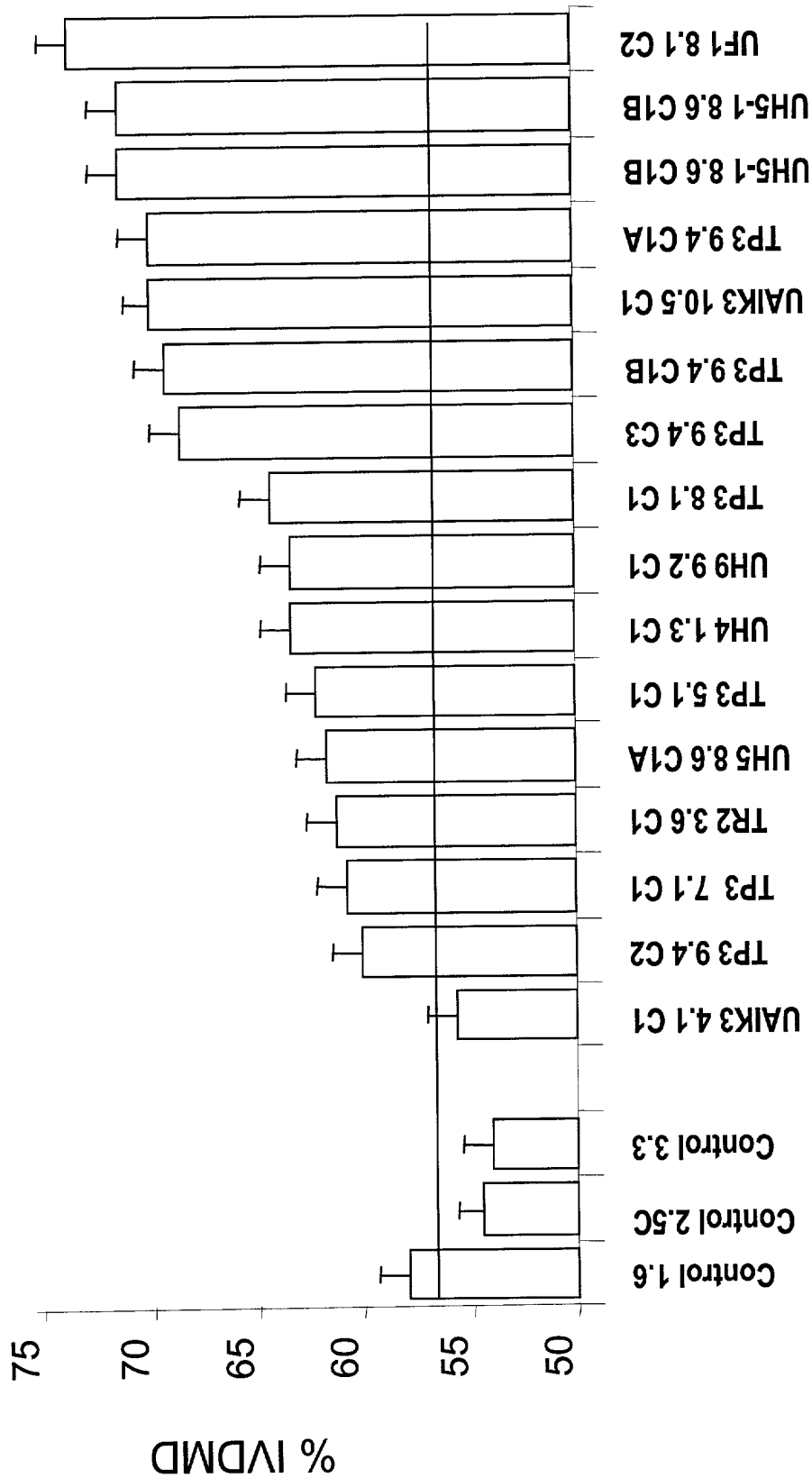
UA1K3* - VAC (Alu-FAE-linker-frameshift)	TP3-VAC (Alu-FAE-linker frameshift)
	TR2-APO (Asp-FAE)
	UH5* -VAC (Alu-FAE-linker-stop)
	UH4* -VAC (Alu-FAE + CaMV-Hyg)
	UH9* - ER (Alu-delNPIR-linker-KDEL)
	UF1* -ER (Asp-FAE-linker -KDEL)

* co-integration vector

Figure 23

In vitro dry mater digestibility of leaf tissue of mature Lolium multiflorum plants expressing FAE under an actin promoter
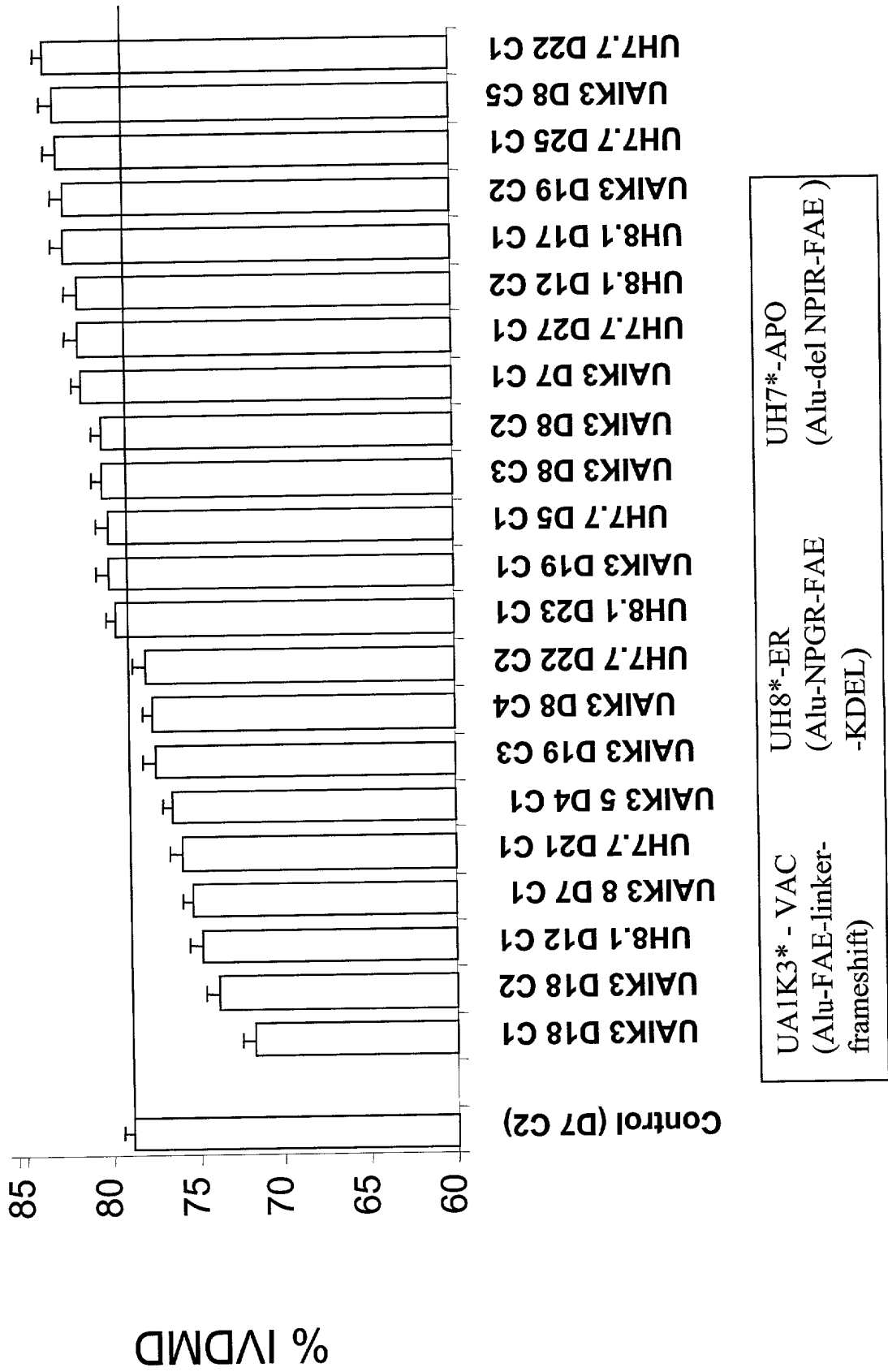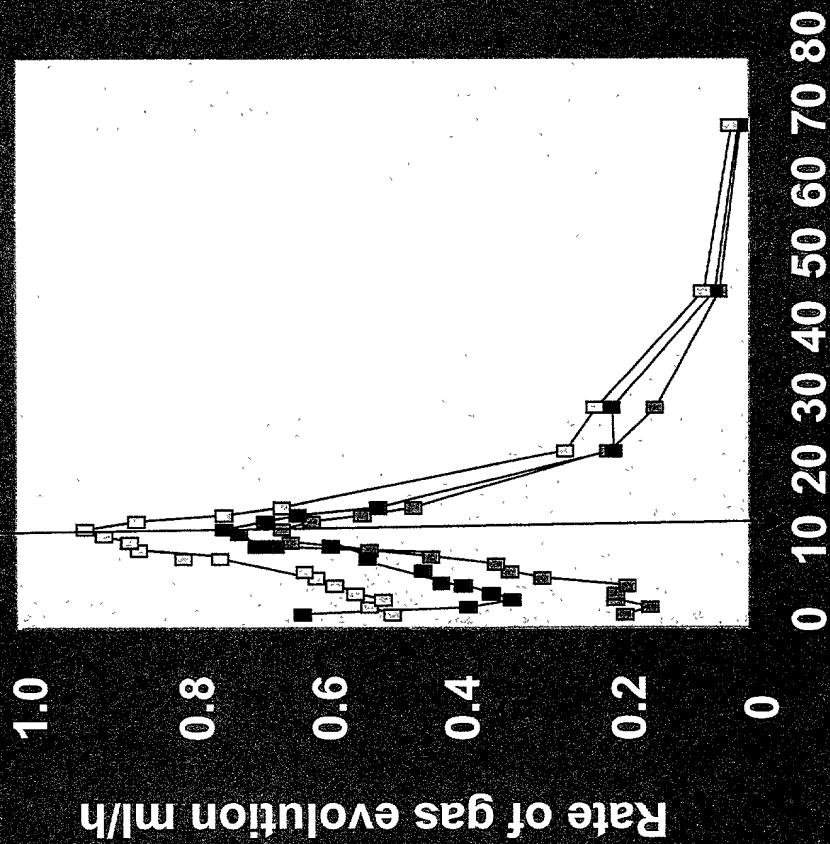

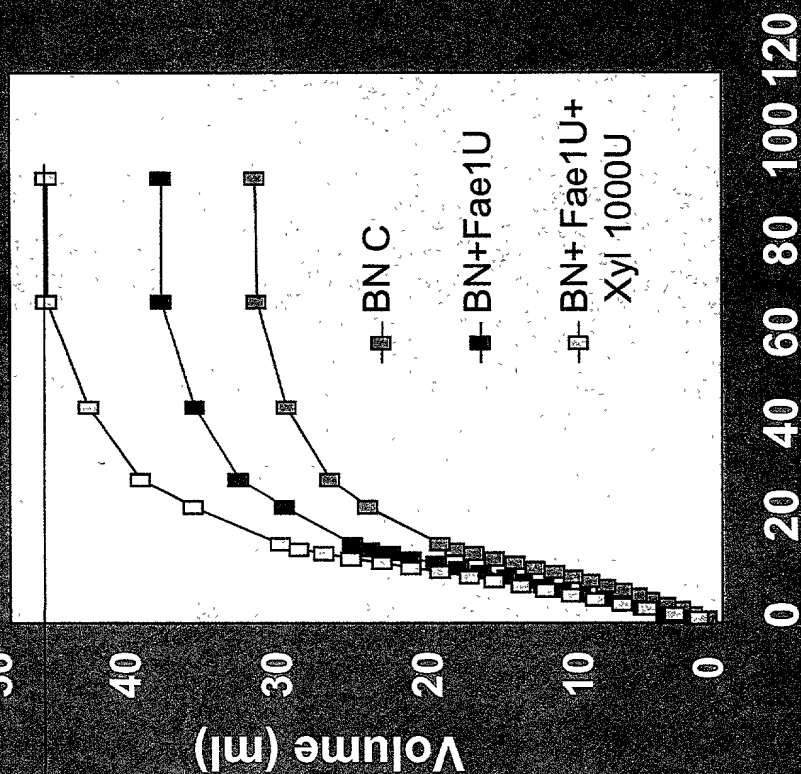| Vector	Description
UA1K3* - VAC (Alu-FAE-linker-frameshift)	UH8*-ER (Alu-NPGR-FAE-KDEL)	UH7*-APO (Alu-del NPIR-FAE)

* co-integration vector

Figure 24

Figure 25

In-vitro fermentation of *Festuca arundinacea* cell walls from cell cultures expressing recombinant FAE1
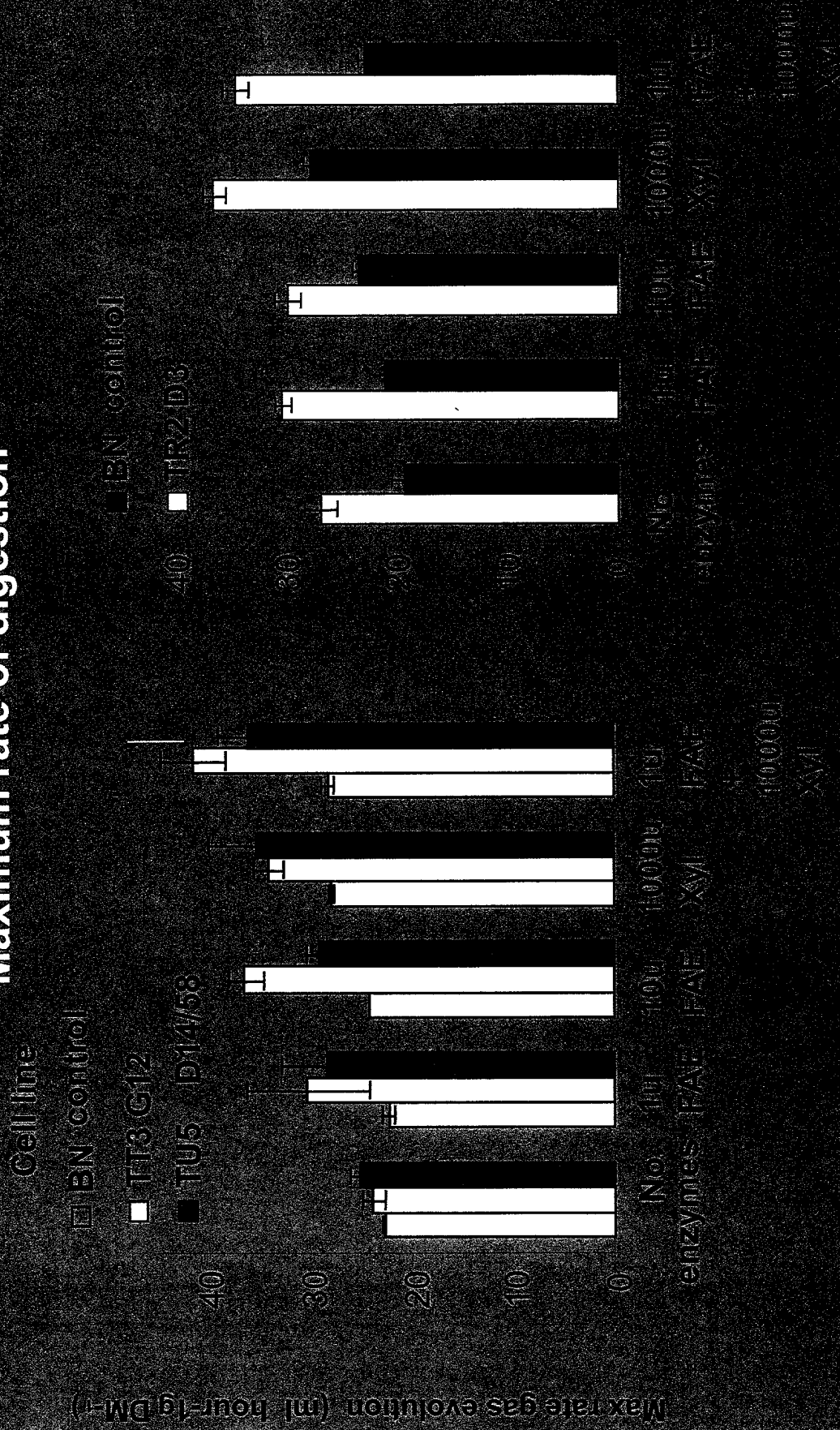
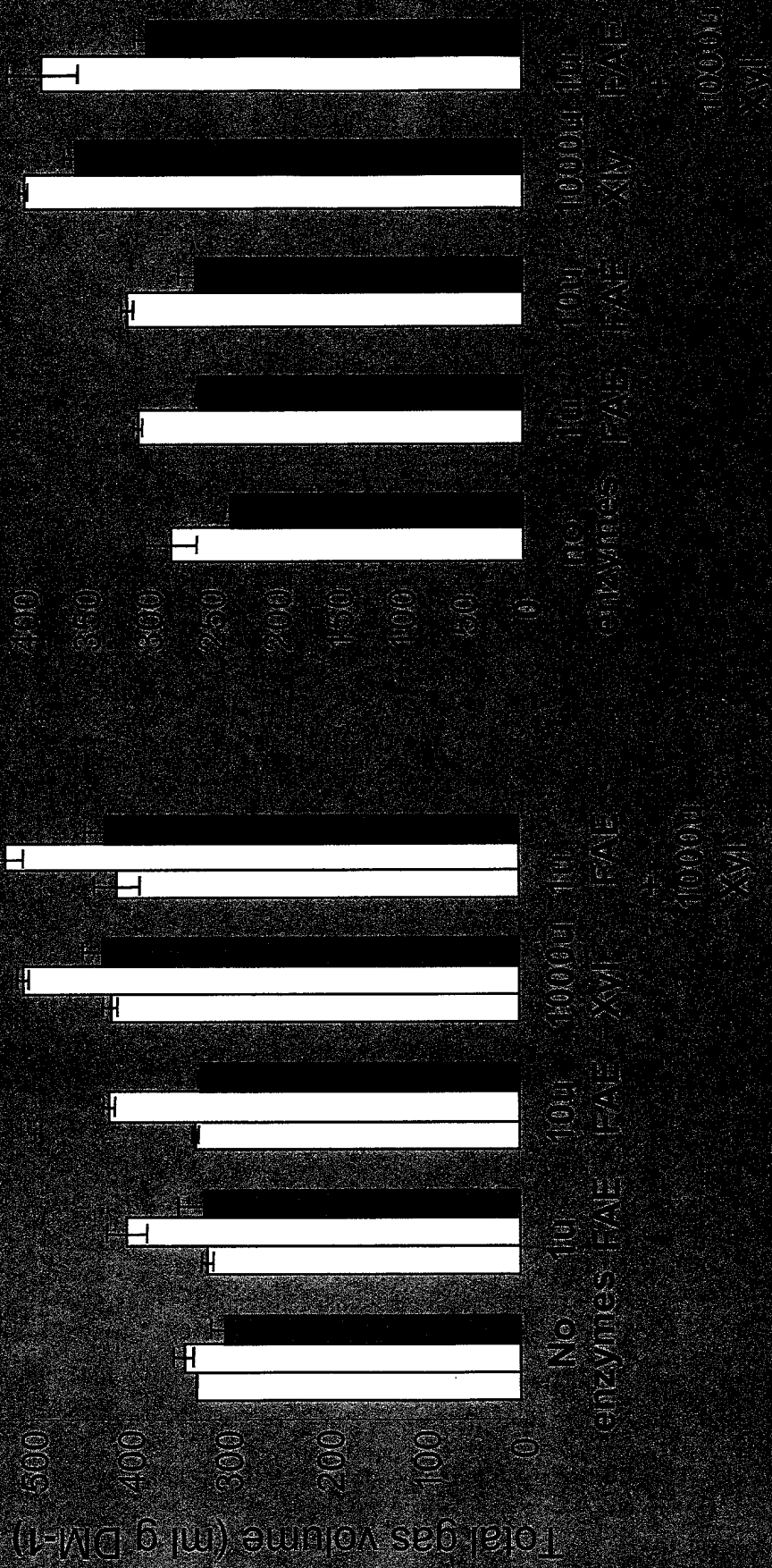## Maximum rate of digestion

Figure 26

Figure 27

Figure 28

Kinetics of FAE activity by ferulic acid release from cell wall under self digestion in *Festuca arundinacea* and stimulation by Xylanase.
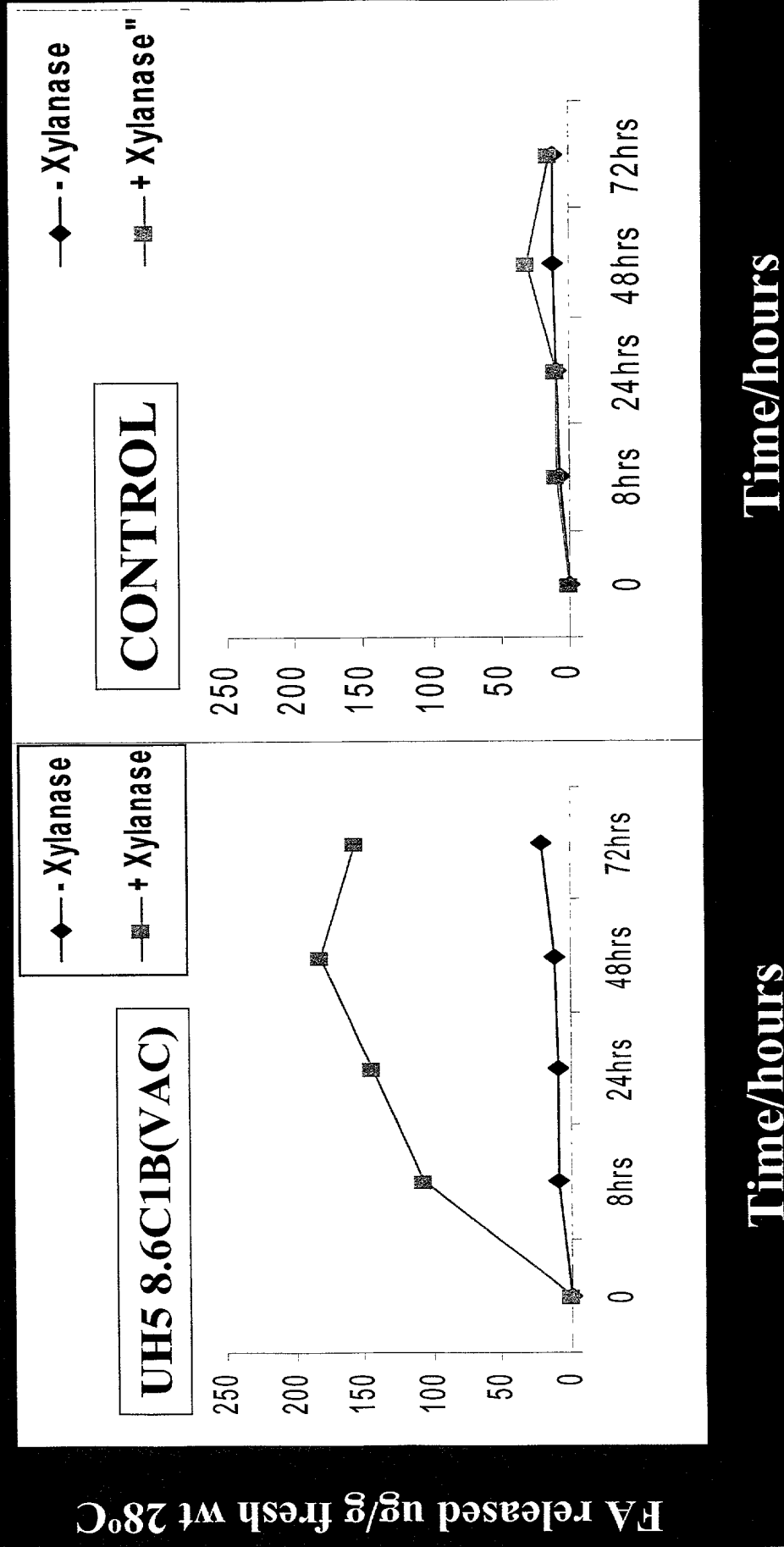

Figure 29

GUS activity under the Lolium See1 senescence promoter in leaves of transgenic plants of *Lolium multiflorum*
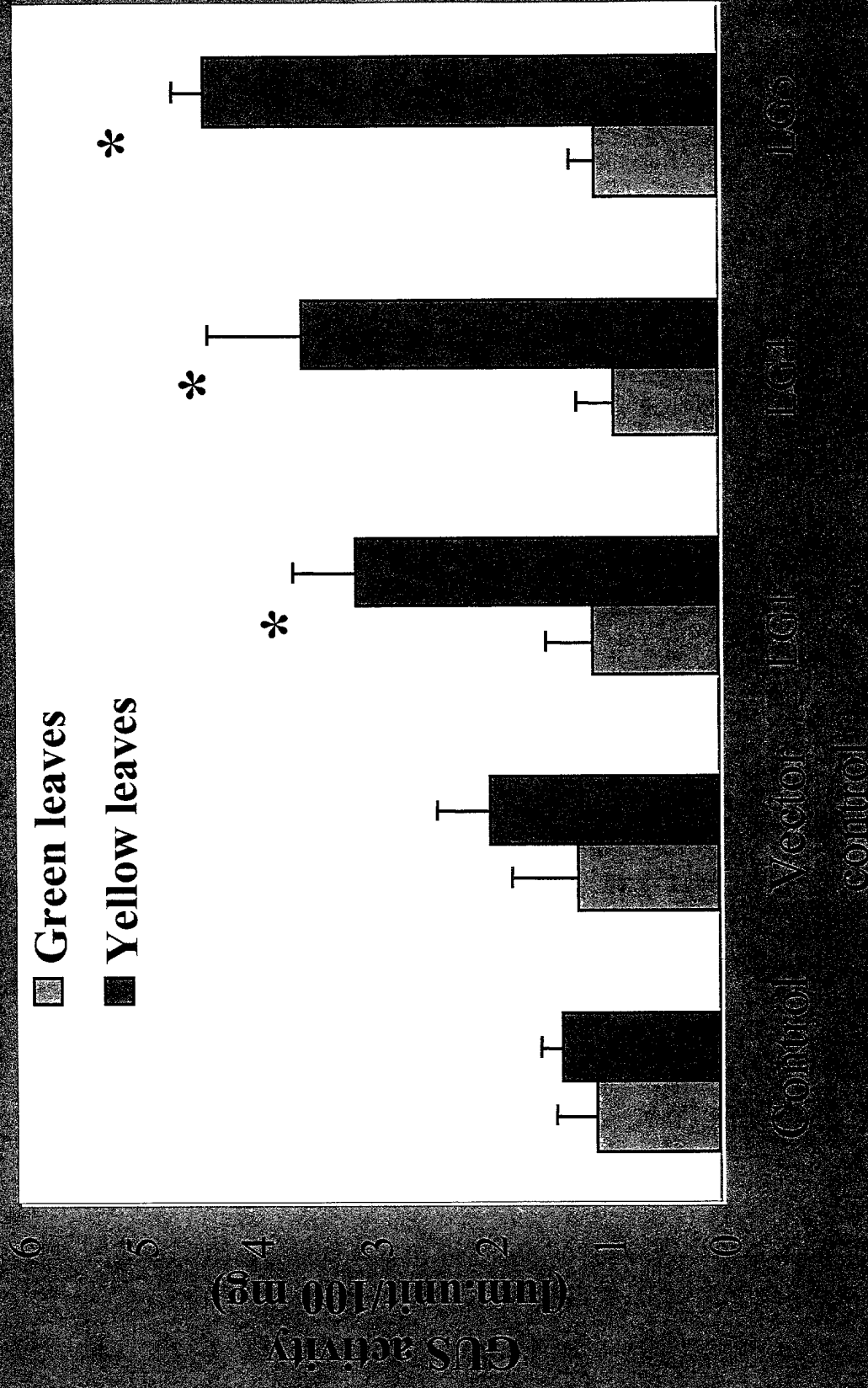
Figure 30

Release of monomeric and dimeric HCAs on self digestion of leaves of vacuolar targeted FAE expressing plants.

Figure 31

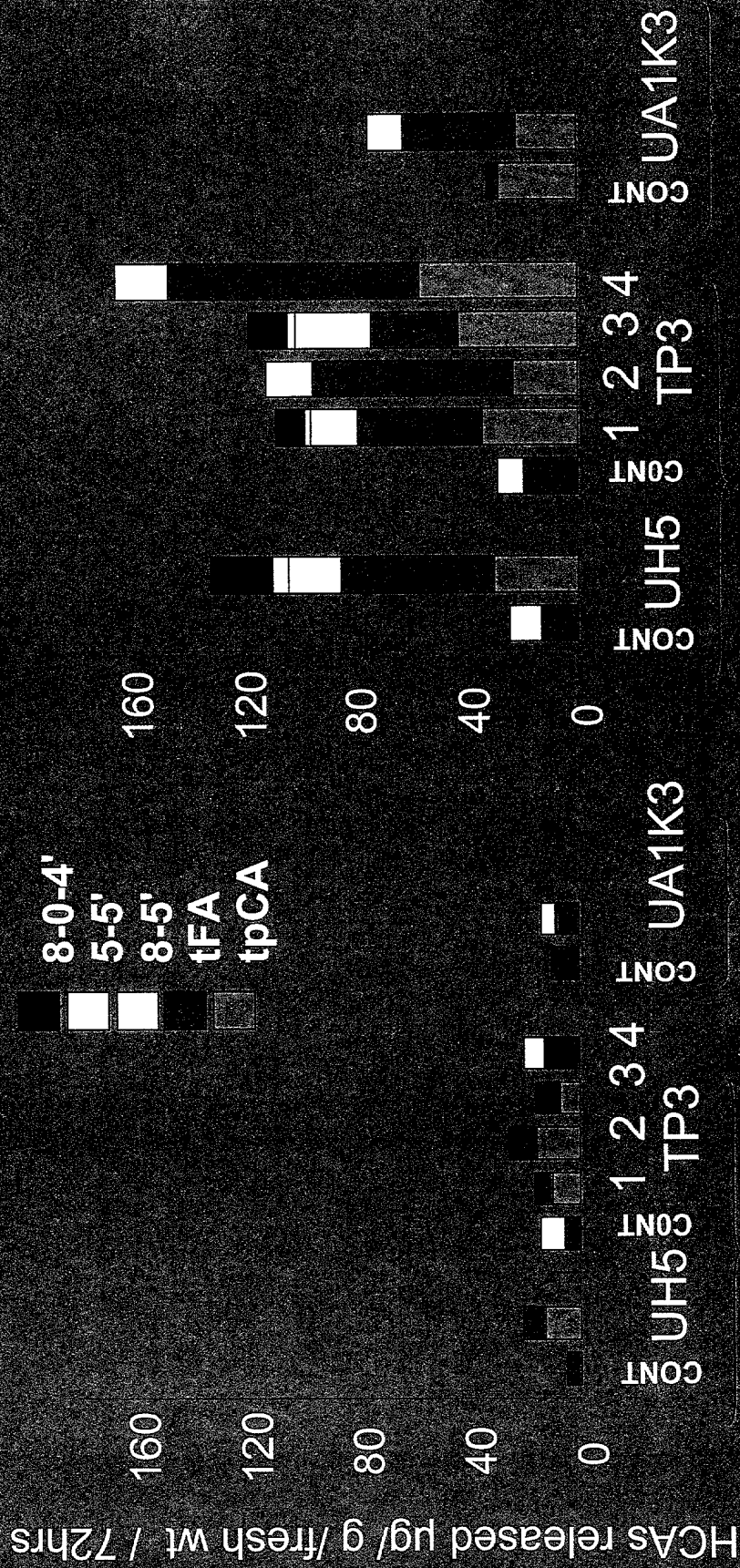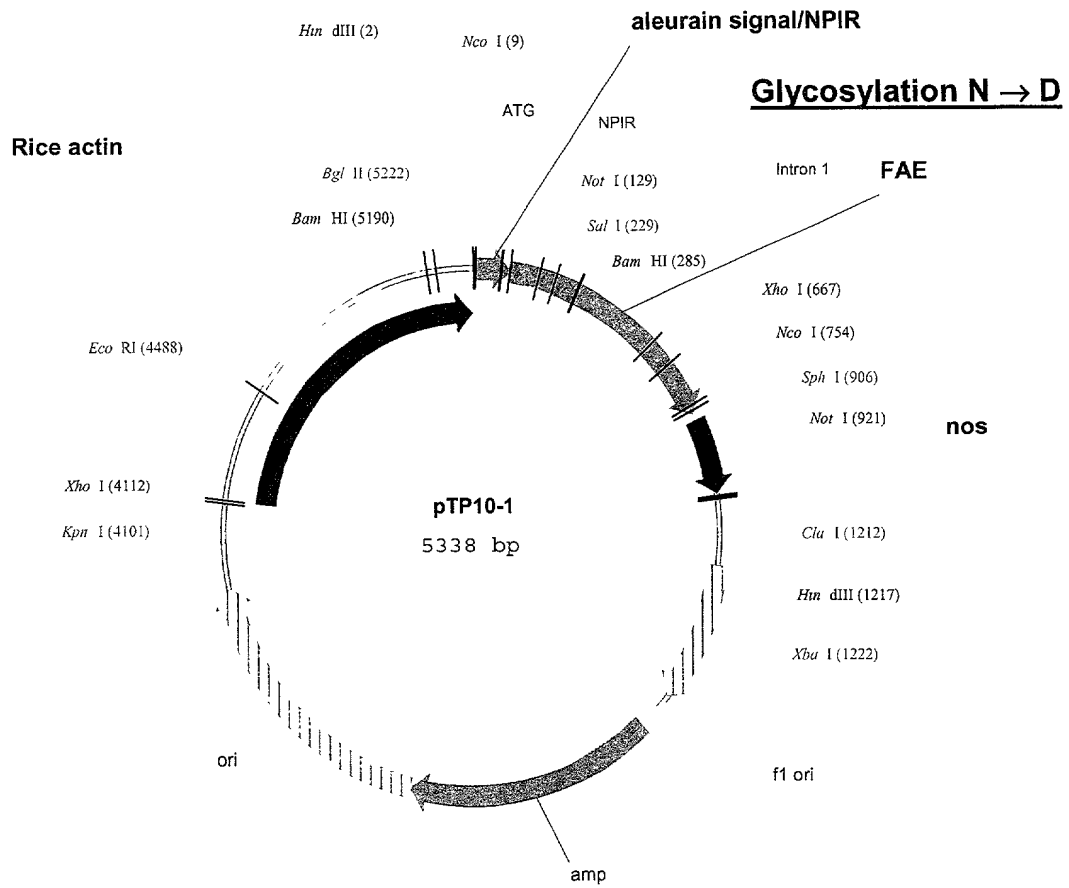Figure 32 A

Figure 32 B

```
                          NcoI
                          ~~~~~~~
                HindIII
                ~~~~~~
                       M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
        1    AAGCTTACCA TGGCCCACGC CGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG

                                             NPIR                            NotI
                                             ~~~~~~~~~~~~~~               ~~~~~~~
                 ·  A   S   S   S   S   F   A   D   S   N   P   I   R   P   V   T   D   R   A   A   A   S   T  ·
       71    TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
                 ·  Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   A   T   I   S   Q   A   A   Y   A
      141    GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC

                 SalI
                 ~~~~~~~
                 D   L   C   N   I   P   S   T   I   I   K   G   E   K   I   Y   N   S   Q   T   D   I   N   G
      211    GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTTACAA TTCTCAAACT GACATTAACG

                 BamHI
                 ~~~~~~
                 ·  W   I   L   R   D   D   S   S   K   E   I   I   T   V   F   R   G   T   G   S   D   T   N  ·
      281    GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACTGGTA GTGATACGAA

                                      Glycosylation
                                      ~~~~
                 ·  L   Q   L   D   T   D   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G   C   E   V
      351    TCTACAACTC GATACTGACT ACACCCTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
                 H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K   Q   Q   V   S
      421    CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA
                 ·  Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A   A   L   T   A  ·
      491    GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
                 ·  A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S   G   N   Q
      561    CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG

                                                    XhoI
                                                    ~~~~~~~
                 A   F   A   S   Y   M   N   D   A   F   Q   A   S   S   P   D   T   T   Q   Y   F   R   V   T
      631    GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA

                                                                           NcoI
                                                                           ~~~~~~~
                 ·  H   A   N   D   G   I   P   N   L   P   P   V   E   Q   G   Y   A   H   G   G   V   E   Y  ·
      701    CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
                 ·  W   S   V   D   P   Y   S   A   Q   N   T   F   V   C   T   G   D   E   V   Q   C   C   E
      771    CTGGAGCGTT GATCCTTACA GCGCCCAGAA CACATTTGTC TGCACTGGGG ATGAAGTGCA GTGCTGTGAG

                                                                               SphI
                                                                               ~~~~
                 A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y   F   G   M   T   S   G   A   C   T   W
      841    GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC GCATGCACCT
```

Figure 32 C

```
            NotI                        KDEL
          ~~~~~~~~~             ~~~~~~~~~~~~~~~
        · P   V   A   A   A   E   T   T   E   G   *
    911 GGCCGGTCGC GGCCGCGGAA ACCACTGAAG GATGAGCTGT AAAGAAGCAG ATCGTTCAAA CATTTGGCAA
    981 TAAAGTTTCT TAAGATTGAA TCCTGTTGCC GGTCTTGCGA TGATTATCAT ATAATTTCTG TTGAATTACG
   1051 TTAAGCATGT AATAATTAAC ATGTAATGCA TGACGTTATT TATGAGATGG GTTTTTATGA TTAGAGTCCC
   1121 GCAATTATAC ATTTAATACG CGATAGAAAA CAAAATATAG CGCGCAAACT AGGATAAATT ATCGCGCGCG

                                   HindIII
                                   ~~~~~~~
                        ClaI        XbaI
                        ~~~~~~      ~~~~~~
   1191 GTGTCATCTA TGTTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC TCCAATTCGC CCTATAGTGA
   1261 GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCTGG CGTTACCCAA
   1331 CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC
   1401 CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGGACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG
   1471 TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC
   1541 CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCGTCAAG CTCTAAATCG GGGGCTCCCT TTAGGGTTCC
   1611 GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAACTTGA TTAGGGTGAT GGTTCACGTA GTGGGCCATC
   1681 GCCCTGATAG ACGGTTTTTC GCCCTTTGAC GTTGGAGTCC ACGTTCTTTA ATAGTGGACT CTTGTTCCAA
   1751 ACTGGAACAA CACTCAACCC TATCTCGGTC TATTCTTTTG ATTTATAAGG GATTTTGCCG ATTTCGGCCT
   1821 ATTGGTTAAA AAATGAGCTG ATTTAACAAA AATTTAACGC GAATTTTAAC AAAATATTAA CGCTTACAAT
   1891 TTAGGTGGCA CTTTTCGGGG AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA
   1961 TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
   2031 TCAACATTTC CGTGTCGCCC TTATTCCCTT TTTTGCGGCA TTTTGCCTTC CTGTTTTTGC TCACCCAGAA
   2101 ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA
   2171 ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG TTTTCCAATG ATGAGCACTT TTAAAGTTCT
   2241 GCTATGTGGC GCGGTATTAT CCCGTATTGA CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT
   2311 CAGAATGACT TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT
   2381 TATGCAGTGC TGCCATAACC ATGAGTGATA ACACTGCGGC CAACTTACTT CTGACAACGA TCGGAGGACC
   2451 GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG GGAACCGGAG
   2521 CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCACGA TGCCTGTAGC AATGGCAACA ACGTTGCGCA
   2591 AACTATTAAC TGGCGAACTA CTTACTCTAG CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA
   2661 AGTTGCAGGA CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT
   2731 GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT
   2801 ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
   2871 TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAACT TCATTTTTAA
   2941 TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT CCCTTAACGT GAGTTTTCGT
   3011 TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT
   3081 CTGCTGCTTG CAAACAAAAA AACCACCGCT ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT
   3151 CTTTTTCCGA AGGTAACTGG CTTCAGCAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG TAGCCGTAGT
   3221 TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
   3291 TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG
   3361 CGGTCGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC GACCTACACC GAACTGAGAT
   3431 ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTAAG
   3501 CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT
   3571 GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA
   3641 AAAACGCCAG CAACGCGGCC TTTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTCC
   3711 TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC TCGCCGCAGC
   3781 CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC AATACGCAAA CCGCCTCTCC
   3851 CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG
   3921 CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT
   3991 ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG GAAACAGCTA TGACCATGAT TACGCCAAGC
```

Figure 32 D

```
                                             KpnI                    XhoI
                                           ~~~~~~~              ~~~~~~~
 4061    GCGCAATTAA CCCTCACTAA AGGGAACAAA AGCTGGGTAC CGGGCCCCCC CTCGAGGTCA TTCATATGCT
 4131    TGAGAAGAGA GTCGGGATAG TCCAAAATAA AACAAAGGTA AGATTACCTG GTCAAAAGTG AAAACATCAG
 4201    TTAAAAGGTG GTATAAGTAA AATATCGGTA ATAAAAGGTG GCCCAAAGTG AAATTTACTC TTTTCTACTA
 4271    TTATAAAAAT TGAGGATGTT TTGTCGGTAC TTTGATACGT CATTTTTGTA TGAATTGGTT TTTAAGTTTA
 4341    TTCGCGATTT GGAAATGCAT ATCTGTATTT GAGTCGGTTT TTAAGTTCGT TGCTTTTGTA AATACAGAGG
 4411    GATTTGTATA AGAAATATCT TTAAAAAACC CATATGCTAA TTTGACATAA TTTTTGAGAA AAATATATAT
           EcoRI
         ~~~~~~~
 4481    TCAGGCGAAT TCCACAATGA ACAATAATAA GATTAAAATA GCTTGCCCCC GTTGCAGCGA TGGGTATTTT
 4551    TTCTAGTAAA ATAAAGATA AACTTAGACT CAAAACATTT ACAAAAACAA CCCCTAAAGT CCTAAAGCCC
 4621    AAAGTGCTAT GCACGATCCA TAGCAAGCCC AGCCCAACCC AACCCAACCC AACCCACCCC AGTGCAGCCA
 4691    ACTGGCAAAT AGTCTCCACC CCCGGCACTA TCACCGTGAG TTGTCCGCAC CACCGCACGT CTCGCAGCCA
 4761    AAAAAAAAAA AAGAAGAAA AAAAGAAAA AGAAAACAG CAGGTGGGTC CGGGTCGTGG GGGCCGGAAA
 4831    AGCGAGGAGG ATCGCGAGCA GCGACGAGGC CCGGCCCTCC CTCCGCTTCC AAAGAAACGC CCCCCATCGC
 4901    CACTATATAC ATACCCCCCC CTCTCCTCCC ATCCCCCAA CCCTACCACC ACCACCACCA CCACCTCCTC
 4971    CCCCCTCGCT GCCGGACGAC GAGCTCCTCC CCCCTCCCCC TCCGCCGCCG CCGGTAACCA CCCCGCCCCT
 5041    CTCCTCTTTC TTTCTCCGTT TTTTTTTTCG TCTCGGTCTC GATCTTTGGC CTTGGTAGTT TGGGTGGGCG
 5111    AGAGCGGCTT CGTCGCCCAG ATCGGTGCGC GGGAGGGGCG GGATCTCGCG GCTGGCGTCT CCGGGCGTGA
           BamHI                                      BglII
         ~~~~~~~                               ~~~~~~~
 5181    GTCGGCCCGG ATCCTCGCGG GGAATGGGGC TCTCGGATGT AGATCTTCTT TCTTTCTTCT TTTTGTGGTA
 5251    GAATTTGAAT CCCTCAGCAT TGTTCATCGG TAGTTTTTCT TTTCATGATT TGTGACAAAT GCAGCCTCGT
 5321    GCGGAGCTTT TTTGTAGC
```
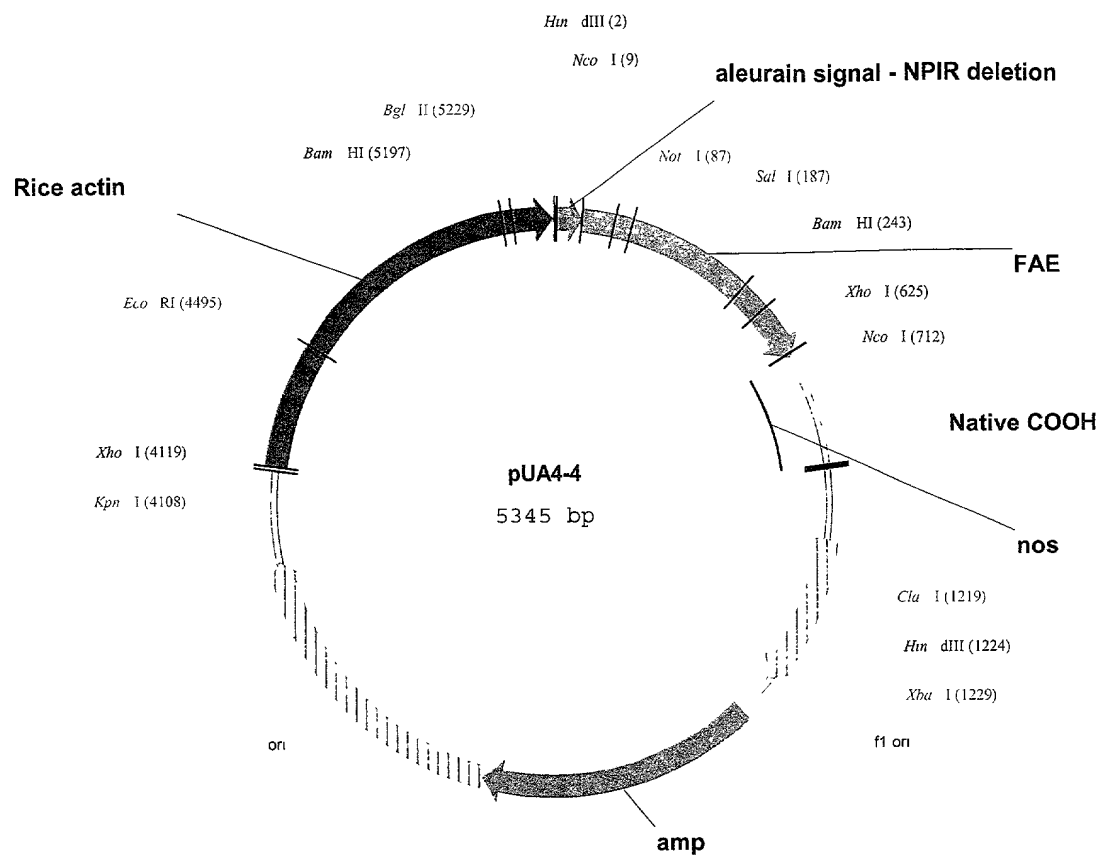
Figure 33 A

Figure 33 B

```
                  NcoI
                  ~~~~~~~

          HindIII
          ~~~~~~

                  M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
      1   AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG
                          NotI
                          ~~~~~~~~~

          ·  A   S   S   R   A   A   A   S   T   Q   G   I   S   E   D   L   Y   S   R   L   V   E   M  ·
     71   TCGCCTCCTC CCGCGCGGCC GCCTCCACGC AGGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT
                                                                      SalI
                                                                      ~~~~~~

          · A   T   I   S   Q   A   A   Y   A   D   L   C   N   I   P   S   T   I   I   K   G   E   K
    141   GGCCACTATC TCCCAAGCTG CCTACGCCGA CCTGTGCAAC ATTCCGTCGA CTATTATCAA GGGAGAGAAA
                                                BamHI
                                                ~~~~~~~

          I   Y   N   S   Q   T   D   I   N   G   W   I   L   R   D   D   S   K   E   I   I   T   V
    211   ATTTACAATT CTCAAACTGA CATTAACGGA TGGATCCTCC GCGACGACAG CAGCAAAGAA ATAATCACCG
          ·  F   R   G   T   G   S   D   T   N   L   Q   L   D   T   N   Y   T   L   T   P   F   D   T  ·
    281   TCTTCCGTGG CACTGGTAGT GATACGAATC TACAACTCGA TACTAACTAC ACCCTCACGC CTTTCGACAC
          ·  L   P   Q   C   N   G   C   E   V   H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q
    351   CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT TATATTGGAT GGGTCTCCGT CCAGGACCAA
          V   E   S   L   V   K   Q   Q   V   S   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L
    421   GTCGAGTCGC TTGTCAAACA GCAGGTTAGC CAGTATCCGG ACTACGCGCT GACCGTGACC GGCCACKCCC
          ·  G   A   S   L   A   A   L   T   A   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T  ·
    491   TCGGCGCCTC CCTGGCGGCA CTCACTGCCG CCCAGCTGTC TGCGACATAC GACAACATCC GCCTGTACAC
                                                                                    XhoI
                                                                                    ~

          ·  F   G   E   P   R   S   G   N   Q   A   F   A   S   Y   M   N   D   A   F   Q   A   S   S
    561   CTTCGGCGAA CCGCGCAGCG GCAATCAGGC CTTCGCGTCG TACATGAACG ATGCCTTCCA AGCCTCGAGC
          P   D   T   T   Q   Y   F   R   V   T   H   A   N   D   G   I   P   N   L   P   P   V   E   Q
    631   CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCCC CCGGTGGAGC
                          NcoI
                          ~~~~~~~

          ·  G   Y   A   H   G   G   V   E   Y   W   S   V   D   P   Y   S   A   Q   N   T   F   V   C  ·
    701   AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCCAGAACA CATTTGTCTG
          · T   G   D   E   V   Q   C   C   E   A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y
    771   CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA CAGGGTGTGA ATAATGCGCA CACGACTTAT
          F   G   M   T   S   G   A   C   T   W   *
    841   TTTGGGATGA CGAGCGGAGC CTGTACATGG TGATCAGTCA TTTCAGCCTC CCCGAGTGTA CCAGGAAAGA
    911   TGGATGTCCT GGAGAGGGGG CCGCGTAACC ACTGAAGGAT GAGCTGTAAA GAAGCAGATC GTTCAAACAT
    981   TTGGCAATAA AGTTTCTTAA GATTGAATCC TGTTGCCGGT CTTGCGATGA TTATCATATA ATTTCTGTTG
   1051   AATTACGTTA AGCATGTAAT AATTAACATG TAATGCATGA CGTTATTTAT GAGATGGGTT TTTATGATTA
   1121   GAGTCCCGCA ATTATACATT TAATACGCGA TAGAAAACAA AATATAGCGC GCAAACTAGG ATAAATTATC
                                                      HindIII
                                                      ~~~~~~~

                                  ClaI          XbaI
                                  ~~~~~~        ~~~~~~
   1191   GCGCGCGGTG TCATCTATGT TACTAGATCG ATAAGCTTCT AGAGCGGCCG GTGGAGCTCC AATTCGCCCT
   1261   ATAGTGAGTC GTATTACGCG CGCTCACTGG CCGTCGTTTT ACAACGTCGT GACTGGGAAA ACCCTGGCGT
   1331   TACCCAACTT AATCGCCTTG CAGCACATCC CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC
   1401   GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGGACGCGCC CTGTAGCGGC GCATTAAGCG
   1471   CGGCGGGTGT GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC
   1541   TTTCTTCCCT TCCTTTCTCG CCACGTTCGC CGGCTTTCCC CGTCAAGCTC TAAATCGGGG GCTCCCTTTA
   1611   GGGTTCCGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTGATGGT TCACGTAGTG
   1681   GGCCATCGCC CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT
   1751   GTTCCAAACT GGAACAACAC TCAACCCTAT CTCGGTCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT
   1821   TCGGCCTATT GGTTAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAACAAA ATATTAACGC
   1891   TTACAATTTA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT GTTTATTTTT CTAAATACAT
   1961   TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA
```
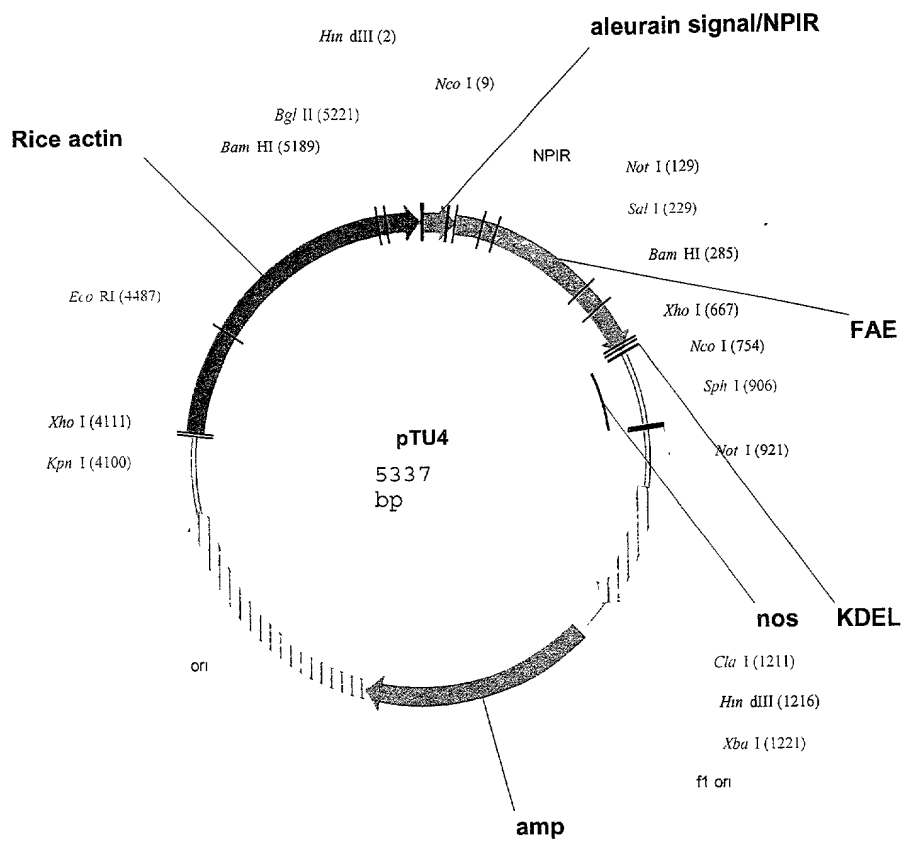
Figure 33 C

```
2031  TGAGTATTCA ACATTTCCGT GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG TTTTTGCTCA
2101  CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG
2171  GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA
2241  AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA
2311  CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA
2381  AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG
2451  GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA
2521  ACCGGAGCTG AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG
2591  TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG
2661  CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG
2731  AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA
2801  GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT
2871  CACTGATTAA GCATTGGTAA CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA
2941  TTTTTAATTT AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG
3011  TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTTTCTGC
3081  GCGTAATCTG CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT
3151  ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA ATACTGTCCT TCTAGTGTAG
3221  CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC
3291  CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA
3361  GGCGCAGCGG TCGGGCTGAA CGGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA
3431  CTGAGATACC TACAGCGTGA GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC
3501  CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAAACGCCT GGTATCTTTA
3571  TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG GGGGCGGAGC
3641  CTATGGAAAA ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG CTGGCCTTTT GCTCACATGT
3711  TCTTTCCTGC GTTATCCCCT GATTCTGTGG ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG
3781  CCGCAGCCGA ACGACCGAGC GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAAACCG
3851  CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA
3921  GTGAGCGCAA CGCAATTAAT GTGAGTTAGC TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC
3991  GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC
                                                        KpnI           XhoI
                                                        ~~~~~~~        ~~~~~~

4061  GCCAAGCGCG CAATTAACCC TCACTAAAGG GAACAAAAGC TGGGTACCGG GCCCCCCCTC GAGGTCATTC
4131  ATATGCTTGA GAAGAGAGTC GGGATAGTCC AAAATAAAAC AAAGGTAAGA TTACCTGGTC AAAAGTGAAA
4201  ACATCAGTTA AAAGGTGGTA TAAGTAAAAT ATCGGTAATA AAAGGTGGCC CAAAGTGAAA TTTACTCTTT
4271  TCTACTATTA TAAAAATTGA GGATGTTTTG TCGGTACTTT GATACGTCAT TTTTGTATGA ATTGGTTTTT
4341  AAGTTTATTC GCGATTTGGA AATGCATATC TGTATTTGAG TCGGTTTTA AGTTCGTTGC TTTTGTAAAT
4411  ACAGAGGGAT TTGTATAAGA AATATCTTTA AAAAACCCAT ATGCTAATTT GACATAATTT TTGAGAAAAA
               EcoRI
               ~~~~~~

4481  TATATATTCA GGCGAATTCC ACAATGAACA ATAATAAGAT TAAAATAGCT TGCCCCCGTT GCAGCGATGG
4551  GTATTTTTTC TAGTAAAAATA AAAGATAAAC TTAGACTCAA AACATTTACA AAAACAACCC CTAAAGTCCT
4621  AAAGCCCAAA GTGCTATGCA CGATCCATAG CAAGCCCAGC CCAACCCAAC CCAACCCAAC CCACCCCAGT
4691  GCAGCCAACT GGCAAATAGT CTCCACCCCC GGCACTATCA CCGTGAGTTG TCCGCACCAC CGCACGTCTC
4761  GCAGCCAAAA AAAAAAAAG AAAGAAAAA AAGAAAAAGA AAAACAGCAG GTGGGTCCGG GTCGTGGGGG
4831  CCGGAAAAGC GAGGAGGATC GCGAGCAGCG ACGAGGCCCG GCCCTCCCTC CGCTTCCAAA GAAACGCCCC
4901  CCATCGCCAC TATATACATA CCCCCCCCTC TCCTCCCATC CCCCCAACCC TACCACCACC ACCACCACCA
4971  CCTCCTCCCC CCTCGCTGCC GGACGACGAG CTCCTCCCCC CTCCCCCTCC GCCGCCGCCG GTAACCACCC
5041  CGCCCCTCTC CTCTTTCTTT CTCCGTTTTT TTTTTCGTCT CGGTCTCGAT CTTTGGCCTT GGTAGTTTGG
5111  GTGGGCGAGA GCGGCTTCGT CGCCCAGATC GGTGCGCGGG AGGGGCGGGA TCTCGCGGCT GGCGTCTCCG
               BamHI                                       BglII
               ~~~~~~~                                     ~~~~~~

5181  GGCGTGAGTC GGCCCGGATC CTCGCGGGGA ATGGGGCTCT CGGATGTAGA TCTTCTTTCT TTCTTCTTTT
5251  TGTGGTAGAA TTTGAATCCC TCAGCATTGT TCATCGGTAG TTTTTCTTTT CATGATTTGT GACAAATGCA
5321  GCCTCGTGCG GAGCTTTTTT GTAGC
```

Figure 34 A

Figure 34 B

```
                  NcoI
                  ~~~~~~~

          HindIII
          ~~~~~~
                  M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
    1     AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG
                                                                                      NotI
                                                                                      ~~~~~~~~
          ·   A   S   S   S   S   F   A   D   S   N   P   I   R   P   V   T   D   R   A   A   A   S   T   ·
   71     TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
          ·   Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   A   T   I   S   Q   A   A   Y   A
  141     GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC
                              SalI
                              ~~~~~~~
          D   L   C   N   I   P   S   T   I   I   K   G   E   K   I   Y   N   S   Q   T   D   I   N   G
  211     GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTTACAA TTCTCAAACT GACATTAACG
          BamHI
          ~~~~~~
          ·   W   I   L   R   D   D   S   S   K   E   I   I   T   V   F   R   G   T   G   S   D   T   N   ·
  281     GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACTGGTA GTGATACGAA
          ·   L   Q   L   D   T   N   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G   C   E   V
  351     TCTACAACTC GATACTAACT ACACCCTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
          H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K   Q   Q   V   S
  421     CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA
          ·   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A   A   L   T   A   ·
  491     GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
          ·   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S   G   N   Q
  561     CGCCCAGCTG TCTGCGACAT CGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
                                                     XhoI
                                                     ~~~~~~~
          A   F   A   S   Y   M   N   D   A   F   Q   A   S   S   P   D   T   T   Q   Y   F   R   V   T
  631     GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA CCCAGATAC GACGCAGTAT TTCCGGGTCA
                                                                                      NcoI
                                                                                      ~~~~~~~
          ·   H   A   N   D   G   I   P   N   L   P   P   V   E   Q   G   Y   A   H   G   G   V   E   Y   ·
  701     CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
          ·   W   S   V   D   P   Y   S   A   Q   N   T   F   V   C   T   G   D   E   V   Q   C   C   E
  771     CTGGAGCGTT GATCCTTACA GCGCCCAGAA CACATTTGTC TGCACTGGGG ATGAAGTGCA GTGCTGTGAG
                                                                                      SphI
                                                                                      ~~~~
          A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y   F   G   M   T   S   G   A   C   T   W
  841     GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC GCATGCACCT
                    NotI
                    ~~~~~~~~~
          ·   P   V   A   A   A   E   P   L   K   D   E   L   *
  911     GGCCGGTCGC GGCCGCGGAA CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT
  981     AAAGTTTCTT AAGATTGAAT CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT TGAATTACGT
 1051     TAAGCATGTA ATAATTAACA TGTAATGCAT GACGTTATTT ATGAGATGGG TTTTTATGAT TAGAGTCCCG
 1121     CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG
                                        HindIII
                                        ~~~~~~~
                        ClaI         XbaI
                        ~~~~~~~      ~~~~~~
 1191     TGTCATCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG
 1261     TCGTATTACG CGCGCTCACT GGCCGTCGTT TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC
 1331     TTAATCGCCT TGCAGCACAT CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
 1401     TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAG CGCGGCGGGT
 1471     GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC
 1541     CTTCCTTTCT CGCCACGTTC GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTTCCG
 1611     ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTGATG GTTCACGTAG TGGGCCATCG
```
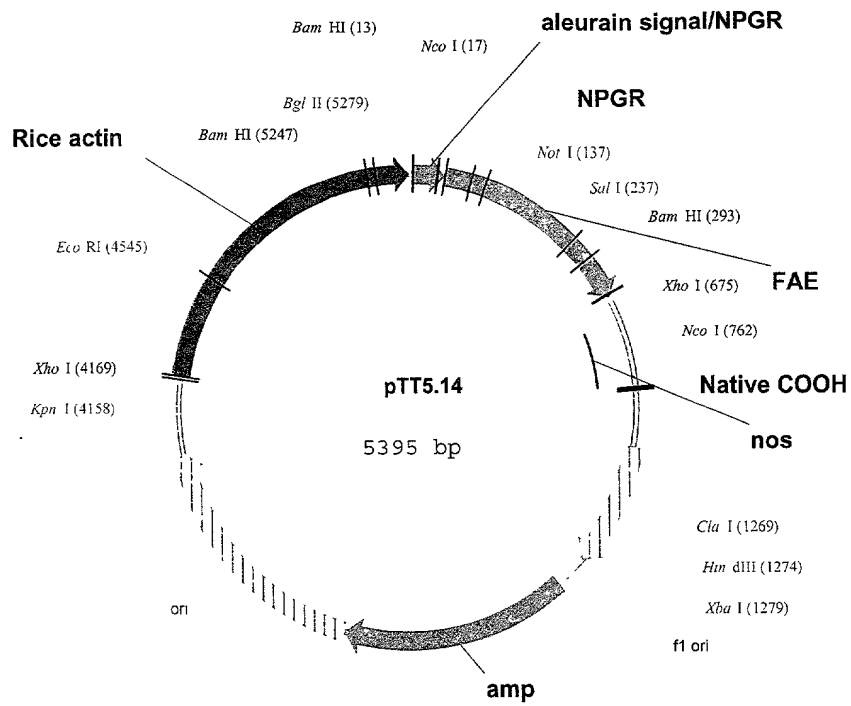
Figure 34 C

```
1681   CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA
1751   CTGGAACAAC ACTCAACCCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGCCGA TTTCGGCCTA
1821   TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG AATTTTAACA AAATATTAAC GCTTACAATT
1891   TAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT
1961   GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT
2031   CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
2101   CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA
2171   CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG
2241   CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
2311   AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT
2381   ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG
2451   AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC
2521   TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA
2591   ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GGCGGATAAA
2661   GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG
2731   AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA
2801   CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT
2871   AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT CATTTTTAAT
2941   TTAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT
3011   CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC
3081   TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
3151   TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT
3221   AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT
3291   GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC
3361   GGTCGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA
3431   CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC
3501   GGCAGGGTCG AACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG
3571   TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA
3641   AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCCT
3711   GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC
3781   GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC CGCCTCTCCC
3851   CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAAGCGGG CAGTGAGCGC
3921   AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA
3991   TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT ACGCCAAGCG
                                                      KpnI             XhoI
                                                      ~~~~~~~          ~~~~~~
4061   CGCAATTAAC CCTCACTAAA GGGAACAAAA GCTGGGTACC GGGCCCCCCC TCGAGGTCAT TCATATGCTT
4131   GAGAAGAGAG TCGGGATAGT CCAAAATAAA ACAAAGGTAA GATTACCTGG TCAAAAGTGA AAACATCAGT
4201   TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCCAAAGTGA AATTTACTCT TTTCTACTAT
4271   TATAAAAATT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTTGTAT GAATTGGTTT TTAAGTTTAT
4341   TCGCGATTTG GAAATGCATA TCTGTATTTG AGTCGGTTTT TAAGTTCGTT GCTTTTGTAA ATACAGAGGG
4411   ATTTGTATAA GAAATATCTT TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA AATATATATT
            EcoRI
            ~~~~~~~
4481   CAGGCGAATT CCACAATGAA CAATAATAAG ATTAAAATAG CTTGCCCCCG TTGCAGCGAT GGGTATTTTT
4551   TCTAGTAAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAAACAAC CCCTAAAGTC CTAAAGCCCA
4621   AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCACCCCA GTGCAGCCAA
4691   CTGGCAAATA GTCTCCACCC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC TCGCAGCCAA
4761   AAAAAAAAAA AGAAAGAAAA AAAAGAAAAA GAAAACAGC AGGTGGGTCC GGGTCGTGGG GGCCGGAAAA
4831   GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCCATCGCC
4901   ACTATATACA TACCCCCCCC TCTCCTCCCA TCCCCCCAAC CCTACCACCA CCACCACCAC CACCTCCTCC
4971   CCCCTCGCTG CCGGACGACG AGCTCCTCCC CCCTCCCCCT CCGCCGCCGC CGGTAACCAC CCCGCCCCTC
5041   TCCTCTTTCT TTCTCCGTTT TTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGGTGGGCGA
5111   GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGTGAG
            BamHI                                        BglII
            ~~~~~~~                                      ~~~~~~
5181   TCGGCCCGGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG
5251   AATTTGAATC CCTCAGCATT GTTCATCGGT AGTTTTTCTT TTCATGATTT GTGACAAATG CAGCCTCGTG
5321   CGGAGCTTTT TTGTAGC
```
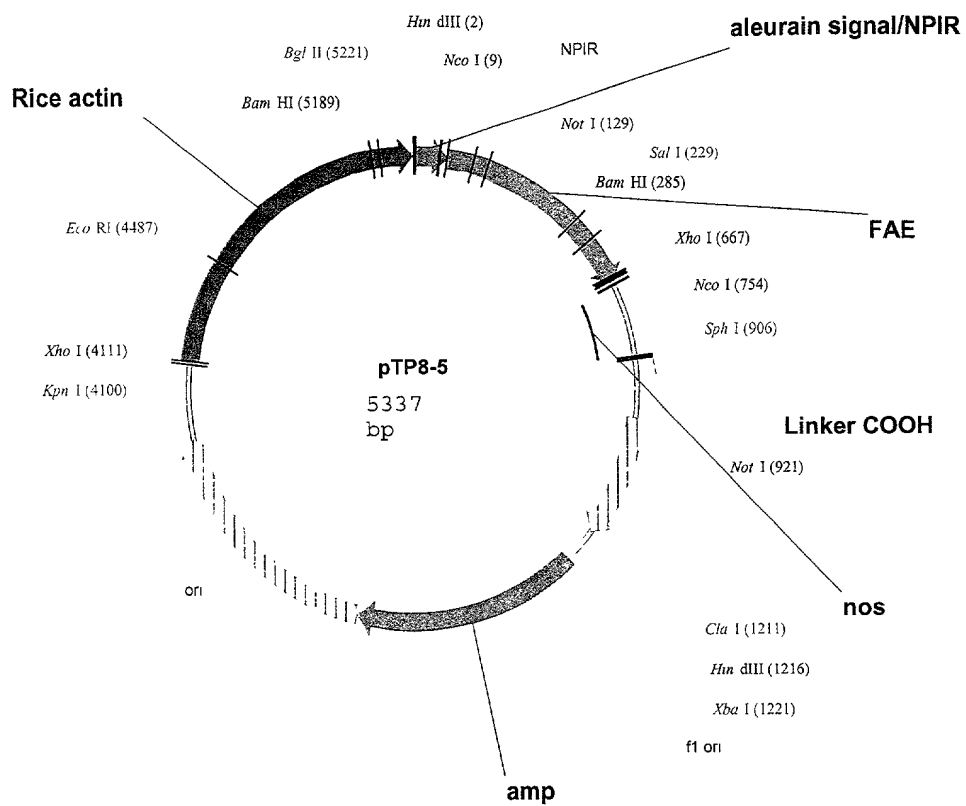
Figure 35 A

Figure 35 B

```
                        NcoI
                       ~~~~~~~
                 BamHI
                ~~~~~~
                            M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A ·
    1   CCTGACGCCG AGGATCCATG GCCCACGCCC GCGTCCTCCT CCTGGCGCTC GCCGTGCTGG CCACGGCCGC
                                                                               NotI
        · V   A   V   A   S   S   S   S   F   A   D   S   N   P   G   R   P   V   T   D   R   A   A
   71   CGTCGCCGTC GCCTCCTCCT CCTCCTTCGC CGACTCCAAC CCGGGCCGGC CCGTCACCGA CCGCGCGGCC
        NotI
        ~~
          A   S   T   Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   A   T   I   S   Q   A   A
  141   GCCTCCACGC AGGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT GGCCACTATC TCCCAAGCTG
                        SalI
                       ~~~~~~~
          ·   Y   A   D   L   C   N   I   P   S   T   I   I   K   G   E   K   I   Y   N   S   Q   T   D ·
  211   CCTACGCCGA CCTGTGCAAC ATTCCGTCGA CTATTATCAA GGGAGAGAAA ATTTACAATT CTCAAACTGA
                      BamHI
                     ~~~~~~
          · I   N   G   W   I   L   R   D   D   S   K   E   I   I   T   V   F   R   G   T   G   S
  281   CATTAACGGA TGGATCCTCC GCGACGACAG CAGCAAAGAA ATAATCACCG TCTTCCGTGG CACTGGTAGT
          D   T   N   L   Q   L   D   T   N   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G   C
  351   GATACGAATC TACAACTCGA TACTAACTAC ACCCTCACGC CTTTCGACAC CCTACCACAA TGCAACGGTT
          · E   V   H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K   Q ·
  421   GTGAAGTACA CGGTGGATAT TATATTGGAT GGGTCTCCGT CCAGGACCAA GTCGAGTCGC TTGTCAAACA
          · Q   V   S   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A   A
  491   GCAGGTTAGC CAGTATCCGG ACTACGCGCT GACCGTGACC GGCCACKCCC TCGGCGCCTC CCTGGCGGCA
          L   T   A   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S   G
  561   CTCACTGCCG CCCAGCTGTC TGCGACATAC GACAACATCC GCCTGTACAC CTTCGGCGAA CCGCGCAGCG
                                                                XhoI
                                                               ~~~~~~~
          · N   Q   A   F   A   S   Y   M   N   D   A   F   Q   A   S   S   P   D   T   T   Q   Y   F ·
  631   GCAATCAGGC CTTCGCGTCG TACATGAACG ATGCCTTCCA AGCCTCGAGC CCAGATACGA CGCAGTATTT
                                                                    NcoI
                                                                   ~~~~
          · R   V   T   H   A   N   D   G   I   P   N   L   P   P   V   E   Q   G   Y   A   H   G   G
  701   CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCCC CCGGTGGAGC AGGGGTACGC CCATGGCGGT
          V   E   Y   W   S   V   D   P   Y   S   A   Q   N   T   F   V   C   T   G   D   E   V   Q   C
  771   GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCCAGAACA CATTTGTCTG CACTGGGGAT GAAGTGCAGT
          · C   E   A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y   F   G   M   T   S   G   A ·
  841   GCTGTGAGGC CCAGGGCGGA CAGGGTGTGA ATAATGCGCA CACGACTTAT TTTGGGATGA CGAGCGGAGC
          · C   T   W   *
  911   CTGTACATGG TGATCAGTCA TTTCAGCCTC CCCGAGTGTA CCAGGAAAGA TGGATGTCCT GGAGAGGGGG
  981   CCGCGTAACC ACTGAAGGAT GAGCTGTAAA GAAGCAGATC GTTCAAACAT TTGGCAATAA AGTTTCTTAA
 1051   GATTGAATCC TGTTGCCGGT CTTGCGATGA TTATCATATA ATTTCTGTTG AATTACGTTA AGCATGTAAT
 1121   AATTAACATG TAATGCATGA CGTTATTTAT GAGATGGGTT TTTATGATTA GAGTCCCGCA ATTATACATT
 1191   TAATACGCGA TAGAAACAA AATATAGCGC GCAAACTAGG ATAAATTATC GCGCGCGGTG TCATCTATGT
                         HindIII
                        ~~~~~~
                 ClaI        XbaI
                ~~~~~~~     ~~~~~~~
 1261   TACTAGATCG ATAAGCTTCT AGAGCGGCCG GTGGAGCTCC AATTCGCCCT ATAGTGAGTC GTATTACGCG
 1331   CGCTCACTGG CCGTCGTTTT ACAACGTCGT GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG
 1401   CAGCACATCC CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT
 1471   GCGCAGCCTG AATGGCGAAT GGGACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG
 1541   CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT TCCTTTCTCG
 1611   CCACGTTCGC CGGCTTTCCC CGTCAAGCTC TAAATCGGGG GCTCCCTTTA GGGTTCCGAT TTAGTGCTTT
 1681   ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTGATGGT TCACGTAGTG GGCCATCGCC CTGATAGACG
 1751   GTTTTTCGCC CTTTGACGTT GGAGTCCACG TTCTTTAATA GTGGACTCTT GTTCCAAACT GGAACAACAC
```

Figure 35 C

```
1821    TCAACCCTAT CTCGGTCTAT TCTTTTGATT TATAAGGGAT TTTGCCGATT TCGGCCTATT GGTTAAAAAA
1891    TGAGCTGATT TAACAAAAAT TTAACGCGAA TTTTAACAAA ATATTAACGC TTACAATTTA GGTGGCACTT
1961    TTCGGGGAAA TGTGCGCGGA ACCCCTATTT GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT
2031    GAGACAATAA CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT
2101    GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG TTTTTGCTCA CCCAGAAACG CTGGTGAAAG
2171    TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
2241    CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT ATGTGGCGCG
2311    GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG
2381    TTGAGTACTC ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC
2451    CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC
2521    GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA
2591    TACCAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC TATTAACTGG
2661    CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA
2731    CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC
2801    GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG
2871    TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA
2941    CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTTAATTT AAAAGGATCT
3011    AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC
3081    AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG CTGCTTGCAA
3151    ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG
3221    TAACTGGCTT CAGCAGAGCG CAGATACCAA ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT
3291    CAAGAACTCT GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC
3361    GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA GGCGCAGCGG TCGGGCTGAA
3431    CGGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
3501    GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC CGGTAAGCGG CAGGGTCGGA
3571    ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC
3641    ACCTCTGACT TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG GGGGCGGAGC CTATGGAAAA ACGCCAGCAA
3711    CGCGGCCTTT TTACGGTTCC TGGCCTTTTG CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCCT
3781    GATTCTGTGG ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
3851    GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC
3921    GATTCATTAA TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT
3991    GTGAGTTAGC TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA
4061    TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCGCG CAATTAACCC
                          KpnI                      XhoI
                          ~~~~~~                    ~~~~~~
4131    TCACTAAAGG GAACAAAAGC TGGGTACCGG GCCCCCCCTC GAGGTCATTC ATATGCTTGA GAAGAGAGTC
4201    GGGATAGTCC AAAATAAAAC AAAGGTAAGA TTACCTGGTC AAAAGTGAAA ACATCAGTTA AAAGGTGGTA
4271    TAAGTAAAAT ATCGGTAATA AAAGGTGGCC CAAAGTGAAA TTTACTCTTT TCTACTATTA TAAAAATTGA
4341    GGATGTTTTG TCGGTACTTT GATACGTCAT TTTTGTATGA ATTGGTTTTT AAGTTTATTC GCGATTTGGA
4411    AATGCATATC TGTATTTGAG TCGGTTTTTA AGTTCGTTGC TTTTGTAAAT ACAGAGGGAT TTGTATAAGA
                                                                           EcoRI
                                                                           ~
4481    AATATCTTTA AAAAACCCAT ATGCTAATTT GACATAATTT TTGAGAAAAA TATATATTCA GGCGAATTCC
4551    ACAATGAACA ATAATAAGAT TAAAATAGCT TGCCCCCGTT GCAGCGATGG GTATTTTTTC TAGTAAAATA
4621    AAAGATAAAC TTAGACTCAA AACATTTACA AAAACAACCC CTAAAGTCCT AAAGCCCAAA GTGCTATGCA
4691    CGATCCATAG CAAGCCCAGC CCAACCCAAC CCAACCCAAC CCACCCCAGT GCAGCCAACT GGCAAATAGT
4761    CTCCACCCCC GGCACTATCA CCGTGAGTTG TCCGCACCAC CGCACGTCTC GCAGCCAAAA AAAAAAAAAG
4831    AAAGAAAAAA AAGAAAAAGA AAAACAGCAG GTGGGTCCGG GTCGTGGGGG CCGGAAAAGC GAGGAGGATC
4901    GCGAGCAGCG ACGAGGCCCG GCCCTCCCTC CGCTTCCAAA GAAACGCCCC CCATCGCCAC TATATACATA
4971    CCCCCCCCTC TCCTCCCATC CCCCCAACCC TACCACCACC ACCACCACCA CCTCCTCCCC CCTCGCTGCC
5041    GGACGACGAG CTCCTCCCCC CTCCCCCTCC GCCGCCGCCG GTAACCACCC CGCCCCTCTC CTCTTTCTTT
5111    CTCCGTTTTT TTTTCGTCT CGGTCTCGAT CTTTGGCCTT GGTAGTTTGG GTGGGCGAGA GCGGCTTCGT
                                                                           BamHI
5181    CGCCCAGATC GGTGCGCGGG AGGGGCGGGA TCTCGCGGCT GGCGTCTCCG GGCGTGAGTC GGCCCGGATC
        BamHI                                 BglII
        ~                                     ~~~~~~~
5251    CTCGCGGGGA ATGGGGCTCT CGGATGTAGA TCTTCTTTCT TTCTTCTTTT TGTGGTAGAA TTTGAATCCC
5321    TCAGCATTGT TCATCGGTAG TTTTTCTTTT CATGATTTGT GACAAATGCA GCCTCGTGCG GAGCTTTTTT
5391    GTAGC
```

Figure 36 A

Figure 36 B

```
                  NcoI
                  ~~~~~~~
          HindIII
          ~~~~~~
                  M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
  1      AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG
                                                                              NotI
                                                                              ~~~~~~~~
         ·   A   S   S   S   S   F   A   D   S   N   P   I   R   P   V   T   D   R   A   A   A   S   T ·
  71     TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
         · Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   A   T   I   S   Q   A   A   Y   A
  141    GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC
                  SalI
                  ~~~~~~~
         D   L   C   N   I   P   S   T   I   I   K   G   E   K   I   Y   N   S   Q   T   D   I   N   G
  211    GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTTACAA TTCTCAAACT GACATTAACG
            BamHI
            ~~~~~~
         ·   W   I   L   R   D   D   S   S   K   E   I   I   T   V   F   R   G   T   G   S   D   T   N ·
  281    GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACTGGTA GTGATACGAA
         · L   Q   L   D   T   N   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G   C   E   V
  351    TCTACAACTC GATACTAACT ACACCCTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
         H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K   Q   Q   V   S
  421    CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA
         ·   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A   A   L   T   A ·
  491    GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
         ·   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S   G   N   Q
  561    CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
                                                              XhoI
                                                              ~~~~~~~
         A   F   A   S   Y   M   N   D   A   F   Q   A   S   S   P   D   T   T   Q   Y   F   R   V   T
  631    GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA
                                                                              NcoI
                                                                              ~~~~~~~
         ·   H   A   N   D   G   I   P   N   L   P   P   V   E   Q   G   Y   A   H   G   G   V   E   Y ·
  701    CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
         · W   S   V   D   P   Y   S   A   Q   N   T   F   V   C   T   G   D   E   V   Q   C   C   E
  771    CTGGAGCGTT GATCCTTACA GCGCCCAGAA CACATTTGTC TGCACTGGGG ATGAAGTGCA GTGCTGTGAG
                                                                              SphI
                                                                              ~~~~
         A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y   F   G   M   T   S   G   A   C   T   W
  841    GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC GCATGCACCT
                  NotI
                  ~~~~~~~~
         · P   V   A   A   A   *
  911    GGCCGGTCGC GGCCGCGTAA CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT
  981    AAAGTTTCTT AAGATTGAAT CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT TGAATTACGT
  1051   TAAGCATGTA ATAATTAACA TGTAATGCAT GACGTTATTT ATGAGATGGG TTTTTATGAT TAGAGTCCCG
  1121   CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG
                                       HindIII
                                       ~~~~~~~
                            ClaI          XbaI
                            ~~~~~~~       ~~~~~~
  1191   TGTCATCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG
  1261   TCGTATTACG CGCGCTCACT GGCCGTCGTT TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC
  1331   TTAATCGCCT TGCAGCACAT CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
  1401   TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAG CGCGGCGGGT
  1471   GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC
  1541   CTTCCTTTCT CGCCACGTTC GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTTCCG
  1611   ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTGATG GTTCACGTAG TGGGCCATCG
```
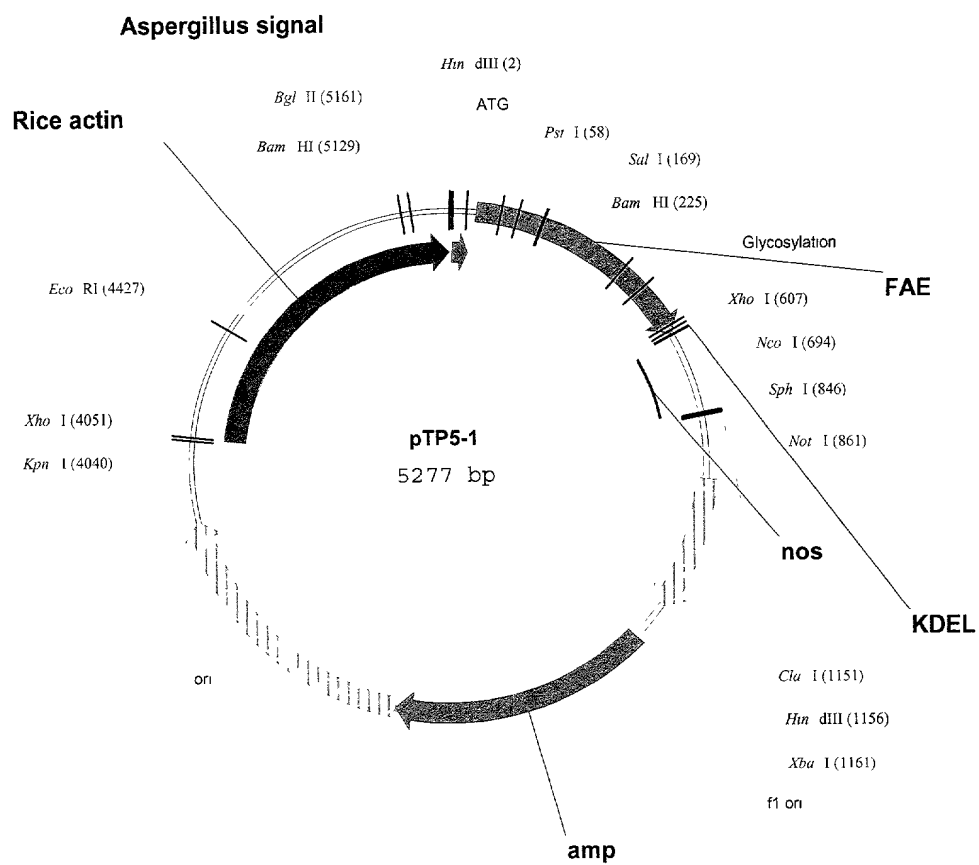
Figure 36 C

```
1681   CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA
1751   CTGGAACAAC ACTCAACCCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGCCGA TTTCGGCCTA
1821   TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG AATTTTAACA AAATATTAAC GCTTACAATT
1891   TAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT
1961   GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT
2031   CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
2101   CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA
2171   CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG
2241   CTATGTGGCG CGGTATTATC CGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC
2311   AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT
2381   ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG
2451   AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC
2521   TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA
2591   ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GGCGGATAAA
2661   GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG
2731   AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA
2801   CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT
2871   AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT CATTTTTAAT
2941   TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT
3011   CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC
3081   TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
3151   TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT
3221   AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT
3291   GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC
3361   GGTCGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA
3431   CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC
3501   GGCAGGGTCG AACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG
3571   TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA
3641   AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCCT
3711   GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC
3781   GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC CGCCTCTCCC
3851   CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAAGCGGG CAGTGAGCGC
3921   AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA
3991   TGTTGTGTGG AATTGTGAGC GGATAACAAT TCACACAGG AAACAGCTAT GACCATGATT ACGCCAAGCG
                                                  KpnI        XhoI
                                                  ~~~~~~      ~~~~~~
4061   CGCAATTAAC CCTCACTAAA GGGAACAAAA GCTGGGTACC GGGCCCCCCC TCGAGGTCAT TCATATGCTT
4131   GAGAAGAGAG TCGGGATAGT CCAAAATAAA ACAAAGGTAA GATTACCTGG TCAAAAGTGA AAACATCAGT
4201   TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCCAAAGTGA AATTTACTCT TTTCTACTAT
4271   TATAAAAATT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTTGTAT GAATTGGTTT TTAAGTTTAT
4341   TCGCGATTTG GAAATGCATA TCTGTATTTG AGTCGGTTTT TAAGTTCGTT GCTTTTGTAA ATACAGAGGG
4411   ATTTGTATAA GAAATATCTT TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA AATATATATT
              EcoRI
              ~~~~~~~
4481   CAGGCGAATT CCACAATGAA CAATAATAAG ATTAAAATAG CTTGCCCCCG TTGCAGCGAT GGGTATTTTT
4551   TCTAGTAAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAAACAAC CCCTAAAGTC CTAAAGCCCA
4621   AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCACCCCA GTGCAGCCAA
4691   CTGGCAAATA GTCTCCACCC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC TCGCAGCCAA
4761   AAAAAAAAAA AGAAAGAAAA AAAAGAAAAA GAAAAACAGC AGGTGGGTCC GGGTCGTGGG GGCCGGAAAA
4831   GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCCATCGCC
4901   ACTATATACA TACCCCCCCC TCTCCTCCCA TCCCCCCAAC CCTACCACCA CCACCACCAC CACCTCCTCC
4971   CCCCTCGCTG CCGGACGACG AGCTCCTCCC CCCTCCCCCT CCGCCGCCGC CGGTAACCAC CCCGCCCCTC
5041   TCCTCTTTCT TTCTCCGTTT TTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGGTGGGCGA
5111   GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGTGAG
              BamHI                               BglII
              ~~~~~~                              ~~~~~~
5181   TCGGCCCGGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG
5251   AATTTGAATC CCTCAGCATT GTTCATCGGT AGTTTTTCTT TTCATGATTT GTGACAAATG CAGCCTCGTG
5321   CGGAGCTTTT TTGTAGC
```

Aspergillus signal

Rice actin

Bgl II (5161)

Bam HI (5129)

Hin dIII (2)

ATG

Pst I (58)

Sal I (169)

Bam HI (225)

Glycosylation

FAE

Xho I (607)

Nco I (694)

Sph I (846)

Not I (861)

Eco RI (4427)

Xho I (4051)

Kpn I (4040)

pTP5-1

5277 bp

nos

KDEL

Cla I (1151)

Hin dIII (1156)

Xba I (1161)

f1 ori

ori

amp

Figure 32 A

Figure 32 B

```
                                                                        PstI
HindIII                                                                 ~~~~~~~
~~~~~~
                M   K   Q   F   S   A   K   H   V   L   A   V   V   V   T   A   G   H   A   L
            A
    1       AAGCTTAACA TGAAGCAGTT CTCCGCCAAA CACGTCCTCG CAGTTGTGGT GACTGCAGGG CACGCCTTAG
            ·   A   S   T   Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   A   T   I   S   Q   A
            ·
   71       CAGCCTCTAC GCAAGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC
                                           SalI
                                           ~~~~~~~
            ·   A   Y   A   D   L   C   N   I   P   S   T   I   I   K   G   E   K   I   Y   N   S   Q   T
  141       TGCCTACGCC GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTTACAA TTCTCAAACT
                       BamHI
                       ~~~~~~~
            D   I   N   G   W   I   L   R   D   D   S   S   K   E   I   I   T   V   F   R   G   T   G
            S
  211       GACATTAACG GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACTGGTA
            ·   D   T   N   L   Q   L   D   T   N   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G
            ·
  281       GTGATACGAA TCTACAACTC GATACTAACT ACACCCTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG
            ·   C   E   V   H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K
  351       TTGTGAAGTA CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA
            Q   Q   V   S   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A
            A
  421       CAGCAGGTTA GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG
            ·   L   T   A   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S
            ·
  491       CACTCACTGC CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG
                                                                    XhoI
                                                                    ~~~~~~~
            ·   G   N   Q   A   F   A   S   Y   M   N   D   A   F   Q   A   S   S   P   D   T   T   Q   Y
  561       CGGCAATCAG GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT
                                                                               NcoI
                                                                               ~~
            F   R   V   T   H   A   N   D   G   I   P   N   L   P   P   V   E   Q   G   Y   A   H   G
            G
  631       TTCCGGGTCA CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG
            ·   V   E   Y   W   S   V   D   P   Y   S   A   Q   N   T   F   V   C   T   G   D   E   V   Q
            ·
  701       GTGTAGAGTA CTGGAGCGTT GATCCTTACA GCGCCCAGAA CACATTTGTC TGCACTGGGG ATGAAGTGCA
            ·   C   C   E   A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y   F   G   M   T   S   G
  771       GTGCTGTGAG GCCCAGGGCG ACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC
            SphI                   NotI
            ~~~~~~                 ~~~~~~~~~
            A   C   T   W   P   V   A   A   A   E   P   L   K   D   E   L   *
  841       GCATGCACCT GGCCGGTCGC GGCCGCGGAA CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC
  911       ATTTGGCAAT AAAGTTTCTT AAGATTGAAT CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT
  981       TGAATTACGT TAAGCATGTA ATAATTAACA TGTAATGCAT GACGTTATTT ATGAGATGGG TTTTTATGAT
 1051       TAGAGTCCCG CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA
                                                             HindIII
                                                             ~~~~~~~
                                       ClaI        XbaI
                                       ~~~~~~      ~~~~~~
 1121       TCGCGCGCGG TGTCATCTAT GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC
 1191       CTATAGTGAG TCGTATTACG CGCGCTCACT GGCCGTCGTT TTACAACGTC GTGACTGGGA AAACCCTGGC
 1261       GTTACCCAAC TTAATCGCCT TGCAGCACAT CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA
 1331       CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAG
 1401       CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC
 1471       GCTTTCTTCC CTTCCTTTCT CGCCACGTTC GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT
```
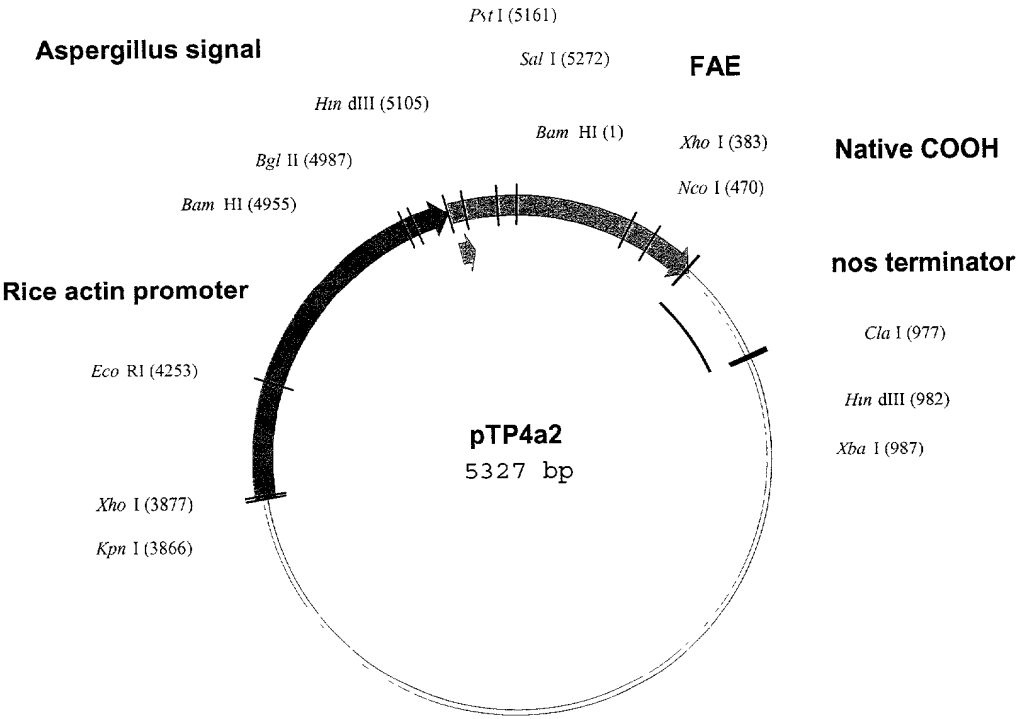
Figure 3 2 C

```
      1541   TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTGATG GTTCACGTAG
      1611   TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC
      1681   TTGTTCCAAA CTGGAACAAC ACTCAACCCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGCCGA
      1751   TTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG AATTTTAACA AAATATTAAC
      1821   GCTTACAATT TAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC
      1891   ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG
      1961   TATGAGTATT CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT
      2031   CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC
      2101   TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT
      2171   TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA
      2241   CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG
      2311   TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT
      2381   CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG
      2451   GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA
      2521   CGTTGCGCAA ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA
      2591   GGCGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT
      2661   GGAGCCGGTG AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG
      2731   TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC
      2801   CTCACTGATT AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT
      2871   CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG
      2941   AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT
      3011   GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG
      3081   CTACCAACTC TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT
      3151   AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT
      3221   ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT
      3291   AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG ACCTACACCG
      3361   AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA
      3431   TCCGGTAAGC GGCAGGGTCG AACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT
      3501   TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
      3571   GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT
      3641   GTTCTTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT
      3711   CGCCGCAGCC GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC
      3781   CGCCTCTCCC CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAAGCGGG
      3851   CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT TAGGCACCAA AGGCTTTACA CTTTATGCTT
      3921   CCGGCTCGTA TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT
                                                                   KpnI             XhoI
                                                                   ~~~~~~~          ~~~~~~~
```

```
      3991   ACGCCAAGCG CGCAATTAAC CCTCACTAAA GGGAACAAAA GCTGGGTACC GGGCCCCCCC TCGAGGTCAT
      4061   TCATATGCTT GAGAAGAGAG TCGGGATAGT CCAAAATAAA ACAAAGGTAA GATTACCTGG TCAAAAGTGA
      4131   AAACATCAGT TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCCAAAGTGA AATTTACTCT
      4201   TTTCTACTAT TATAAAAATT GAGGATGTTT TGTCGGTACT TTGATACGTC ATTTTTGTAT GAATTGGTTT
      4271   TTAAGTTTAT TCGCGATTTG GAAATGCATA TCTGTATTTG AGTCGGTTTT TAAGTTCGTT GCTTTTGTAA
      4341   ATACAGAGGG ATTTGTATAA GAAATATCTT TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA
                          EcoRI
                          ~~~~~~~
```

```
      4411   AATATATATT CAGGCGAATT CCACAATGAA CAATAATAAG ATTAAAATAG CTTGCCCCCG TTGCAGCGAT
      4481   GGGTATTTTT TCTAGTAAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAAACAAC CCCTAAAGTC
      4551   CTAAAGCCCA AAGTGCTATG CACGATCCAT AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCACCCCA
      4621   GTGCAGCCAA CTGGCAAATA GTCTCCACCC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC
      4691   TCGCAGCCAA AAAAAAAAAA AGAAAGAAAA AAAAGAAAAA GAAAACAGC AGGTGGGTCC GGGTCGTGGG
      4761   GGCCGGAAAA GCGAGGAGGA TCGCGAGCAG CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC
      4831   CCCCATCGCC ACTATATACA TACCCCCCCC TCTCCTCCCA TCCCCCCAAC CCTACCACCA CCACCACCAC
      4901   CACCTCCTCC CCCCTCGCTG CCGGACGACG AGCTCCTCCC CCCTCCCCCT CCGCCGCCGC CGGTAACCAC
      4971   CCCGCCCCTC TCCTCTTTCT TTCTCCGTTT TTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT
      5041   GGGTGGGCGA GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC
                          BamHI                                           BglII
                          ~~~~~~~                                         ~~~~~~
```

```
      5111   CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT
      5181   TTTGTGGTAG AATTTGAATC CCTCAGCATT GTTCATCGGT AGTTTTTCTT TTCATGATTT GTGACAAATG
      5251   CAGCCTCGTG CGGAGCTTTT TTGTAGC
```

Figure 38 A

Aspergillus signal

Pst I (5161)

Sal I (5272)

FAE

Hin dIII (5105)

Bam HI (1)

Xho I (383)

Native COOH

Bgl II (4987)

Nco I (470)

Bam HI (4955)

Rice actin promoter

nos terminator

Eco RI (4253)

Cla I (977)

pTP4a2
5327 bp

Hin dIII (982)

Xba I (987)

Xho I (3877)

Kpn I (3866)

Figure 38 B

```
                                    BamHI
                                    ~~~~~
       · I   L   R   D   D   S   S   K   E   I   I   T   V   F   R   G   T   G   S   D   T   N   L
   1     GATCCTCCGC GACGACAGCA GCAAAGAAAT AATCACCGTC TTCCGTGGCA CTGGTAGTGA TACGAATCTA
         Q   L   D   T   N   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G   C   E   V   H   G
  71     CAACTCGATA CTAACTACAC CCTCACGACT TTCGACACCC TACCACAATG CAACGGTTGT GAAGTACACG
       · G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K   Q   Q   V   S   Q   ·
 141     GTGGATATTA TATTGGATGG GTCTCCGTCC AGGACCAAGT CGAGTCGCTT GTCAAACAGC AGGTTAGCCA
       · Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A   A   L   T   A   A
 211     GTATCCGGAC TACGCGCTGA CCGTGACCGG CCACKCCCTC GGCGCCTCCC TGGCGGCACT CACTGCCGCC
         Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S   G   N   Q   A   F
 281     CAGCTGTCTG CGACATACGA CAACATCCGC CTGTACACCT TCGGCGAACC GCGCAGCGGC AATCAGGCCT
                                                                     XhoI
                                                                     ~~~~~~~
       · A   S   Y   M   N   D   A   F   Q   A   S   S   P   D   T   T   Q   Y   F   R   V   T   H   ·
 351     TCGCGTCGTA CATGAACGAT GCCTTCCAAG CCTCGAGCCC AGATACGACG CAGTATTTCC GGGTCACTCA
                                                                     NcoI
                                                                     ~~~~~~
       · A   N   D   G   I   P   N   L   P   P   V   E   Q   G   Y   A   H   G   G   V   E   Y   W
 421     TGCCAACGAC GGCATCCCAA ACCTGCCCCC GGTGGAGCAG GGGTACGCCC ATGGCGGTGT AGAGTACTGG
         S   V   D   P   Y   S   A   Q   N   T   F   V   C   T   G   D   E   V   Q   C   C   E   A   Q
 491     AGCGTTGATC CTTACAGCGC CCAGAACACA TTTGTCTGCA CTGGGGATGA AGTGCAGTGC TGTGAGGCCC
       · G   G   Q   G   V   N   N   A   H   T   T   Y   F   G   M   T   S   G   A   C   T   W   *   ·
 561     AGGGCGGACA GGGTGTGAAT AATGCGCACA CGACTTATTT TGGGATGACG AGCGGAGCCT GTACATGGTG
                                                    · *
 631     ATCAGTCATT TCAGCCTCCC CGAGTGTACC AGGAAAGATG GATGTCCTGG AGAGGGGGCC GCGTAACCAC
 701     TGAAGGATGA GCTGTAAAGA AGCAGATCGT TCAAACATTT GGCAATAAAG TTTCTTAAGA TTGAATCCTG
 771     TTGCCGGTCT TGCGATGATT ATCATATAAT TTCTGTTGAA TTACGTTAAG CATGTAATAA TTAACATGTA
 841     ATGCATGACG TTATTTATGA GATGGGTTTT TATGATTAGA GTCCCGCAAT TATACATTTA ATACGCGATA
                                                                                      ClaI
 911     GAAACAAAA TATAGCGCGC AAACTAGGAT AAATTATCGC GCGCGGTGTC ATCTATGTTA CTAGATCGAT
                       XbaI
                       ~~~~~~~
                                    HindIII
                                    ~~~~~~
 981     AAGCTTCTAG AGCGGCCGGT GGAGCTCCAA TTCGCCCTAT AGTGAGTCGT ATTACGCGCG CTCACTGGCC
1051     GTCGTTTTAC AACGTCGTGA CTGGGAAAAC CCTGGCGTTA CCCAACTTAA TCGCCTTGCA GCACATCCCC
1121     CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG CCCGCACCGA TCGCCCTTCC CAACAGTTGC GCAGCCTGAA
1191     TGGCGAATGG GACGCGCCCT GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG TGGTTACGCG CAGCGTGACC
1261     GCTACACTTG CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC CTTTCTCGCC ACGTTCGCCG
1331     GCTTTCCCCG TCAAGCTCTA AATCGGGGGC TCCCTTTAGG GTTCCGATTT AGTGCTTTAC GGCACCTCGA
1401     CCCCAAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT
1471     TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT TCCAAACTGG AACAACACTC AACCCTATCT
1541     CGGTCTATTC TTTTGATTTA TAAGGGATTT TGCCGATTTC GGCCTATTGG TTAAAAAATG AGCTGATTTA
1611     ACAAAAATTT AACGCGAATT TTAACAAAAT ATTAACGCTT ACAATTAGG TGGCACTTTT CGGGGAAATG
1681     TGCGCGGAAC CCCTATTTGT TTATTTTTCT AAATACATTC AAATATGTAT CCGCTCATGA GACAATAACC
1751     CTGATAAATG CTTCAATAAT ATTGAAAAAG GAAGAGTATG AGTATTCAAC ATTTCCGTGT CGCCCTTATT
1821     CCCTTTTTTG CGGCATTTTG CCTTCCTGTT TTTGCTCACC CAGAAACGCT GGTGAAAGTA AAAGATGCTG
1891     AAGATCAGTT GGGTGCACGA GTGGGTTACA TCGAACTGGA TCTCAACAGC GGTAAGATCC TTGAGAGTTT
1961     TCGCCCCGAA GAACGTTTTC CAATGATGAG CACTTTTAAA GTTCTGCTAT GTGGCGCGGT ATTATCCCGT
2031     ATTGACGCCG GCAAGAGCA ACTCGGTCGC CGCATACACT ATTCTCAGAA TGACTTGGTT GAGTACTCAC
2101     CAGTCACAGA AAAGCATCTT ACGGATGGCA TGACAGTAAG AGAATTATGC AGTGCTGCCA TAACCATGAG
2171     TGATAACACT GCGGCCAACT TACTTCTGAC AACGATCGGA GGACCGAAGG AGCTAACCGC TTTTTTGCAC
2241     AACATGGGGG ATCATGTAAC TCGCCTTGAT CGTTGGGAAC CGGAGCTGAA TGAAGCCATA CCAAACGACG
2311     AGCGTGACAC CACGATGCCT GTAGCAATGG CAACAACGTT GCGCAAACTA TTAACTGGCG AACTACTTAC
2381     TCTAGCTTCC CGGCAACAAT TAATAGACTG GATGGAGGCG GATAAAGTTG CAGGACCACT TCTGCGCTCG
2451     GCCCTTCCGG CTGGCTGGTT TATTGCTGAT AAATCTGGAG CCGGTGAGCG TGGGTCTCGC GGTATCATTG
2521     CAGCACTGGG GCCAGATGGT AAGCCCTCCC GTATCGTAGT TATCTACACG ACGGGGAGTC AGGCAACTAT
2591     GGATGAACGA AATAGACAGA TCGCTGAGAT AGGTGCCTCA CTGATTAAGC ATTGGTAACT GTCAGACCAA
2661     GTTTACTCAT ATATACTTTA GATTGATTTA AAACTTCATT TTTAATTTAA AAGGATCTAG GTGAAGATCC
```

Figure 38 C

```
2731   TTTTTGATAA TCTCATGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG ACCCCGTAGA
2801   AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT GCTTGCAAAC AAAAAAACCA
2871   CCGCTACCAG CGGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACTCTTTT TCCGAAGGTA ACTGGCTTCA
2941   GCAGAGCGCA GATACCAAAT ACTGTCCTTC TAGTGTAGCC GTAGTTAGGC CACCACTTCA AGAACTCTGT
3011   AGCACCGCCT ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT
3081   CTTACCGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG GGGGGTTCGT
3151   GCACACAGCC CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA CAGCGTGAGC TATGAGAAAG
3221   CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA GGGTCGGAAC AGGAGAGCGC
3291   ACGAGGGAGC TTCCAGGGGG AAACGCCTGG TATCTTTATA GTCCTGTCGG GTTTCGCCAC CTCTGACTTG
3361   AGCGTCGATT TTTGTGATGC TCGTCAGGGG GGCGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTTT
3431   ACGGTTCCTG GCCTTTTGCT GGCCTTTTGC TCACATGTTC TTTCCTGCGT TATCCCCTGA TTCTGTGGAT
3501   AACCGTATTA CCGCCTTTGA GTGAGCTGAT ACCGCTCGCC GCAGCCGAAC GACCGAGCGC AGCGAGTCAG
3571   TGAGCGAGGA AGCGGAAGAG CGCCCAATAC GCAAACCGCC TCTCCCCGCG CGTTGGCCGA TTCATTAATG
3641   CAGCTGGCAC GACAGGTTTC CCGACTGGAA AGCGGGCAGT GAGCGCAACG CAATTAATGT GAGTTAGCTC
3711   ACTCATTAGG CACCCCAGGC TTTACACTTT ATGCTTCCGG CTCGTATGTT GTGTGGAATT GTGAGCGGAT
3781   AACAATTTCA CACAGGAAAC AGCTATGACC ATGATTACGC CAAGCGCGCA ATTAACCCTC ACTAAAGGGA
                                          KpnI             XhoI
                                          ~~~~~~           ~~~~~~
3851   ACAAAGCTG GGTACCGGGC CCCCCCTCGA GGTCATTCAT ATGCTTGAGA GAGAGTCGG GATAGTCCAA
3921   AATAAAACAA AGGTAAGATT ACCTGGTCAA AAGTGAAAAC ATCAGTTAAA AGGTGGTATA AGTAAAATAT
3991   CGGTAATAAA AGGTGGCCCA AAGTGAAATT TACTCTTTTC TACTATTATA AAAATTGAGG ATGTTTTGTC
4061   GGTACTTTGA TACGTCATTT TTGTATGAAT TGGTTTTTAA GTTTATTCGC GATTTGGAAA TGCATATCTG
4131   TATTTGAGTC GGTTTTTAAG TTCGTTGCTT TTGTAAATAC AGAGGGATTT GTATAAGAAA TATCTTTAAA
                                                                        EcoRI
                                                                        ~~~~~~~
4201   AAACCCATAT GCTAATTTGA CATAATTTTT GAGAAAAATA TATATTCAGG CGAATTCCAC AATGAACAAT
4271   AATAAGATTA AAATAGCTTG CCCCCGTTGC AGCGATGGGT ATTTTTTCTA GTAAAATAAA AGATAAACTT
4341   AGACTCAAAA CATTTACAAA AACAACCCCT AAAGTCCTAA AGCCCAAAGT GCTATGCACG ATCCATAGCA
4411   AGCCCAGCCC AACCCAACCC AACCCAACCC ACCCCAGTGC AGCCAACTGG CAAATAGTCT CCACCCCCGG
4481   CACTATCACC GTGAGTTGTC CGCACCACCG CACGTCTCGC AGCCAAAAAA AAAAAAAGAA AGAAAAAAAA
4551   GAAAAAGAAA AACAGCAGGT GGGTCCGGGT CGTGGGGGCC GGAAAAGCGA GGAGGATCGC GAGCAGCGAC
4621   GAGGCCCGGC CCTCCCTCCG CTTCCAAAGA AACGCCCCCC ATCGCCACTA TATACATACC CCCCCCTCTC
4691   CTCCCATCCC CCCAACCCTA CCACCACCAC CACCACCACC TCCTCCCCCC TCGCTGCCGG ACGACCGAGCT
4761   CCTCCCCCCT CCCCCTCCGC CGCCGCCGGT AACCACCCCG CCCCTCTCCT CTTTCTTTCT CCGTTTTTTT
4831   TTTCGTCTCG GTCTCGATCT TTGGCCTTGG TAGTTTGGGT GGGCGAGAGC GGCTTCGTCG CCCAGATCGG
                                                                        BamHI
                                                                        ~~~~~~~
4901   TGCGCGGGAG GGGCGGGATC TCGCGGCTGG CGTCTCCGGG CGTGAGTCGG CCCGGATCCT CGCGGGGAAT
                                                            BglII
                                                            ~~~~~~~
4971   GGGGCTCTCG GATGTAGATC TTCTTTCTTT CTTCTTTTTG TGGTAGAATT TGAATCCCTC AGCATTGTTC
                                                                        HindIII
                                                                        ~
5041   ATCGGTAGTT TTTCTTTTCA TGATTTGTGA CAAATGCAGC CTCGTGCGGA GCTTTTTTGT AGCAAGCTTA
                                                                        PstI
                                                                        ~~~~~~~
          M  K  Q   F  S  A   K  H  V  L   A  V  V   V  T  A   G  H  A  L   A  A  S ·
5111   ACATGAAGCA GTTCTCCGCC AAACACGTCC TCGCAGTTGT GGTGACTGCA GGGCACGCCT TAGCAGCCTC
          · T  Q  G   I´ S  E  D   L  Y  S   R  L  V   E  M  A  T   I  S  Q   A  A  Y
5181   TACGCAAGGC ATCTCCGAAG ACCTCTACAG CCGTTTAGTC GAAATGGCCA CTATCTCCCA AGCTGCCTAC
                                                        SalI
                                                        ~~~~~~
          A  D  L  C   N  I  P   S  T  I   I  K  G  E   K  I  Y   N  S  Q   T  D  I  N
5251   GCCGACCTGT GCAACATTCC GTCGACTATT ATCAAGGGAG AGAAAATTTA CAATTCTCAA ACTGACATTA
                                          B
                                          ~
                             ·  G  W
5321                         ACGGATG
```
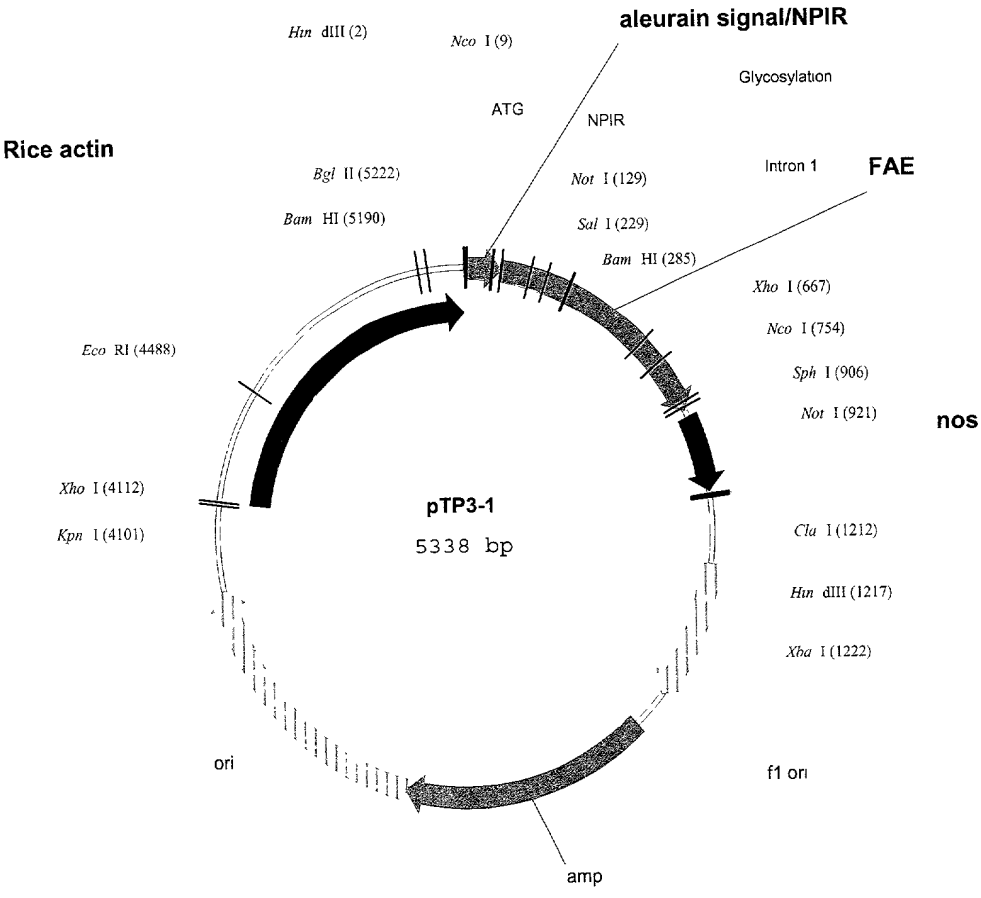
Figure 39 A

Figure 39 A

Figure 39 B

```
            NcoI
            ~~~~~~~
    HindIII
    ~~~~~~

            M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
   1    AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG

                                            NPIR                        NotI
                                            ~~~~~~~~~~~~~~              ~~~~~~~~
        ·   A   S   S   S   S   F   A   D   S   N   P   I   R   P   V   T   D   R   A   A   A   S   T  ·
  71    TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
        ·   Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   A   T   I   S   Q   A   A   Y   A
 141    GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC

                SalI
                ~~~~~~~
        D   L   C   N   I   P   S   T   I   I   K   G   E   K   I   Y   N   S   Q   T   D   I   N   G
 211    GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTTACAA TTCTCAAACT GACATTAACG

            BamHI
            ~~~~~~
        ·   W   I   L   R   D   D   S   S   K   E   I   I   T   V   F   R   G   T   G   S   D   T   N  ·
 281    GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACTGGTA GTGATACGAA

                Glycosylation
                ~~~~
        ·   L   Q   L   D   T   N   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G   C   E   V
 351    TCTACAACTC GATACTAACT ACACCCTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
          H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K   Q   Q   V   S
 421    CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA
        ·   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A   A   L   T   A  ·
 491    GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
        ·   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S   G   N   Q
 561    CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG

                                            XhoI
                                            ~~~~~~~
        A   F   A   S   Y   M   N   D   A   F   Q   A   S   S   P   D   T   T   Q   Y   F   R   V   T
 631    GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA

                                                                    NcoI
                                                                    ~~~~~~~
        ·   H   A   N   D   G   I   P   N   L   P   P   V   E   Q   G   Y   A   H   G   G   V   E   Y  ·
 701    CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
        ·   W   S   V   D   P   Y   S   A   Q   N   T   F   V   C   T   G   D   E   V   Q   C   C   E
 771    CTGGAGCGTT GATCCTTACA GCGCCCAGAA CACATTTGTC TGCACTGGGG ATGAAGTGCA GTGCTGTGAG

                                                                            SphI
                                                                            ~~~~
        A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y   F   G   M   T   S   G   A   C   T   W
 841    GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC GCATGCACCT
```

Figure 39 C

```
            NotI                     KDEL
         ~~~~~~~~          ~~~~~~~~~~~~~
        · P  V  A   A  A  E   T   T  E  G    *
  911   GGCCGGTCGC GGCCGCGGAA ACCACTGAAG GATGAGCTGT AAAGAAGCAG ATCGTTCAAA CATTTGGCAA
  981   TAAAGTTTCT TAAGATTGAA TCCTGTTGCC GGTCTTGCGA TGATTATCAT ATAATTTCTG TTGAATTACG
 1051   TTAAGCATGT AATAATTAAC ATGTAATGCA TGACGTTATT TATGAGATGG GTTTTTATGA TTAGAGTCCC
 1121   GCAATTATAC ATTTAATACG CGATAGAAAA CAAAATATAG CGCGCAAACT AGGATAAATT ATCGCGCGCG

                               HindIII
                               ~~~~~~~
                        ClaI       XbaI
                        ~~~~~~     ~~~~~~
 1191   GTGTCATCTA TGTTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC TCCAATTCGC CCTATAGTGA
 1261   GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCTGG CGTTACCCAA
 1331   CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC ACCGATCGCC
 1401   CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGGACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG
 1471   TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC
 1541   CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCGTCAAG CTCTAAATCG GGGGCTCCCT TTAGGGTTCC
 1611   GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAACTTGA TTAGGGTGAT GGTTCACGTA GTGGGCCATC
 1681   GCCCTGATAG ACGGTTTTTC GCCCTTTGAC GTTGGAGTCC ACGTTCTTTA ATAGTGGACT CTTGTTCCAA
 1751   ACTGGAACAA CACTCAACCC TATCTCGGTC TATTCTTTTG ATTTATAAGG GATTTTGCCG ATTTCGGCCT
 1821   ATTGGTTAAA AAATGAGCTG ATTTAACAAA AATTTAACGC GAATTTTAAC AAAATATTAA CGCTTACAAT
 1891   TTAGGTGGCA CTTTTCGGGG AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA
 1961   TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
 2031   TCAACATTTC CGTGTCGCCC TTATTCCCTT TTTTGCGGCA TTTTGCCTTC CTGTTTTTGC TCACCCAGAA
 2101   ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA
 2171   ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG TTTTCCAATG ATGAGCACTT TTAAAGTTCT
 2241   GCTATGTGGC GCGGTATTAT CCCGTATTGA CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT
 2311   CAGAATGACT TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAT
 2381   TATGCAGTGC TGCCATAACC ATGAGTGATA ACACTGCGGC CAACTTACTT CTGACAACGA TCGGAGGACC
 2451   GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG GGAACCGGAG
 2521   CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCACGA TGCCTGTAGC AATGGCAACA ACGTTGCGCA
 2591   AACTATTAAC TGGCGAACTA CTTACTCTAG CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA
 2661   AGTTGCAGGA CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT
 2731   GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT
 2801   ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
 2871   TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAACT TCATTTTTAA
 2941   TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT CCCTTAACGT GAGTTTTCGT
 3011   TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT
 3081   CTGCTGCTTG CAAACAAAAA AACCACCGCT ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT
 3151   CTTTTTCCGA AGGTAACTGG CTTCAGCAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG TAGCCGTAGT
 3221   TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC
 3291   TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA TAAGGCGCAG
 3361   CGGTCGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC GACCTACACC GAACTGAGAT
 3431   ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG CGGACAGGT ATCCGGTAAG
 3501   CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT
 3571   GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA
 3641   AAAACGCCAG CAACGCGGCC TTTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTCC
 3711   TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC TCGCCGCAGC
 3781   CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC AATACGCAAA CCGCCTCTCC
 3851   CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG
 3921   CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT
```

Figure 39 D

```
3991   ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG GAAACAGCTA TGACCATGAT TACGCCAAGC
                                                  KpnI                 XhoI
                                                  ~~~~~~~              ~~~~~~~
4061   GCGCAATTAA CCCTCACTAA AGGGAACAAA AGCTGGGTAC CGGGCCCCCC CTCGAGGTCA TTCATATGCT
4131   TGAGAAGAGA GTCGGGATAG TCCAAAATAA AACAAAGGTA AGATTACCTG GTCAAAAGTG AAAACATCAG
4201   TTAAAAGGTG GTATAAGTAA AATATCGGTA ATAAAAGGTG GCCCAAAGTG AAATTTACTC TTTTCTACTA
4271   TTATAAAAAT TGAGGATGTT TTGTCGGTAC TTTGATACGT CATTTTTGTA TGAATTGGTT TTTAAGTTTA
4341   TTCGCGATTT GGAAATGCAT ATCTGTATTT GAGTCGGTTT TTAAGTTCGT TGCTTTTGTA AATACAGAGG
4411   GATTTGTATA AGAAATATCT TTAAAAAACC CATATGCTAA TTTGACATAA TTTTTGAGAA AAATATATAT
               EcoRI
               ~~~~~~~
4481   TCAGGCGAAT TCCACAATGA ACAATAATAA GATTAAAATA GCTTGCCCCC GTTGCAGCGA TGGGTATTTT
4551   TTCTAGTAAA ATAAAGATA  AACTTAGACT CAAAACATTT ACAAAAACAA CCCCTAAAGT CCTAAAGCCC
4621   AAAGTGCTAT GCACGATCCA TAGCAAGCCC AGCCCAACCC AACCCAACCC AACCCACCCC AGTGCAGCCA
4691   ACTGGCAAAT AGTCTCCACC CCCGGCACTA TCACCGTGAG TTGTCCGCAC CACCGCACGT CTCGCAGCCA
4761   AAAAAAAAAA AAGAAGAAA  AAAAAGAAAA AGAAAAACAG CAGGTGGGTC CGGGTCGTGG GGGCCGGAAA
4831   AGCGAGGAGG ATCGCGAGCA GCGACGAGGC CCGGCCCTCC CTCCGCTTCC AAAGAAACGC CCCCCATCGC
4901   CACTATATAC ATACCCCCCC CTCTCCTCCC ATCCCCCCAA CCCTACCACC ACCACCACCA CCACCTCCTC
4971   CCCCCTCGCT GCCGGACGAC GAGCTCCTCC CCCCTCCCCC TCCGCCGCCG CCGGTAACCA CCCCGCCCCT
5041   CTCCTCTTTC TTTCTCCGTT TTTTTTTTCG TCTCGGTCTC GATCTTTGGC CTTGGTAGTT TGGGTGGGCG
5111   AGAGCGGCTT CGTCGCCCAG ATCGGTGCGC GGGAGGGGCG GGATCTCGCG GCTGGCGTCT CCGGGCGTGA
               BamHI                              BglII
               ~~~~~~~                            ~~~~~~~
5181   GTCGGCCCGG ATCCTCGCGG GGAATGGGGC TCTCGGATGT AGATCTTCTT TCTTTCTTCT TTTTGTGGTA
5251   GAATTTGAAT CCCTCAGCAT TGTTCATCGG TAGTTTTTCT TTTCATGATT TGTGACAAAT GCAGCCTCGT
5321   GCGGAGCTTT TTTGTAGC
```

HindIII (2)

NcoI (9)

ATG

aleurain_signal/NPIR

NPIR

BamHI (5189)

AvaI (103)

SmaI (105)

Rice_actin

NotI (129)

EcoRI (4487)

BamHI (285)

Glycosylation

Intron_1

AvaI (4111)

FAE

pTU5
5337 bp

PvuII (568)

AvaI (667)

PvuII (3878)

NcoI (754)

NotI (921)

Redundancy (stop)

ApaLI (3384)

KDEL

ori

nos

ClaI (1211)

HindIII (1216)

XbaI (1221)

PvuII (1365)

f1_ori

ApaLI (2138)

amp

Figure 40 A

Sequence for pTU5

```
     HindIII NcoI
     ~~~~~~ ~~~~~~
   1 AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT
     TTCGAATGGT ACCGGGTGCG GGCGCAGGAG GAGGACCGCG AGCGGCACGA

  51 GGCCACGGCC GCCGTCGCCG TCGCCTCCTC CTCCTCCTTC GCCGACTCCA
     CCGGTGCCGG CGGCAGCGGC AGCGGAGGAG GAGGAGGAAG CGGCTGAGGT


     SmaI
     ~~~~~~
     AvaI                              NotI
     ~~~~~~                          ~~~~~~~~~
 101 ACCCGGGCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC GCAGGGCATC
     TGGGCCCGGC CGGGCAGTGG CTGGCGCGCC GGCGGAGGTG CGTCCCGTAG

 151 TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC
     AGGCTTCTGG AGATGTCGGC AAATCAGCTT TACCGGTGAT AGAGGGTTCG

 201 TGCCTACGCC GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA
     ACGGATGCGG CTGGACACGT TGTAAGGCAG CTGATAATAG TTCCCTCTCT


                                         BamHI
                                       ~~~~~~
 251 AAATTTACAA TTCTCAAACT GACATTAACG GATGGATCCT CCGCGACGAC
     TTTAAATGTT AAGAGTTTGA CTGTAATTGC CTACCTAGGA GGCGCTGCTG

 301 AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACTGGTA GTGATACGAA
     TCGTCGTTTC TTTATTAGTG GCAGAAGGCA CCGTGACCAT CACTATGCTT

 351 TCTACAACTC GATACTAACT ACACCCTCAC GCCTTTCGAC ACCCTACCAC
     AGATGTTGAG CTATGATTGA TGTGGGAGTG CGGAAAGCTG TGGGATGGTG

 401 AATGCAACGG TTGTGAAGTA CACGGTGGAT ATTATATTGG ATGGGTCTCC
     TTACGTTGCC AACACTTCAT GTGCCACCTA TAATATAACC TACCCAGAGG

 451 GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA GCCAGTATCC
     CAGGTCCTGG TTCAGCTCAG CGAACAGTTT GTCGTCCAAT CGGTCATAGG

 501 GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG
     CCTGATGCGC GACTGGCACT GGCCGGTGMG GGAGCCGCGG AGGGACCGCC


                       PvuII
                     ~~~~~~
 551 CACTCACTGC CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC
     GTGAGTGACG GCGGGTCGAC AGACGCTGTA TGCTGTTGTA GGCGGACATG

 601 ACCTTCGGCG AACCGCGCAG CGGCAATCAG GCCTTCGCGT CGTACATGAA
     TGGAAGCCGC TTGGCGCGTC GCCGTTAGTC CGGAAGCGCA GCATGTACTT


                  AvaI
```

Fig. 40 B

```
                          ~~~~~~~
  651  CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA
       GCTACGGAAG GTTCGGAGCT CGGGTCTATG CTGCGTCATA AAGGCCCAGT

  701  CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC
       GAGTACGGTT GCTGCCGTAG GGTTTGGACG GGGGCCACCT CGTCCCCATG

            NcoI
            ~~~~~~
  751  GCCCATGGCG GTGTAGAGTA CTGGAGCGTT GATCCTTACA GCGCCCAGAA
       CGGGTACCGC CACATCTCAT GACCTCGCAA CTAGGAATGT CGCGGGTCTT

  801  CACATTTGTC TGCACTGGGG ATGAAGTGCA GTGCTGTGAG GCCCAGGGCG
       GTGTAAACAG ACGTGACCCC TACTTCACGT CACGACACTC CGGGTCCCGC

  851  GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGC
       CTGTCCCACA CTTATTACGC GTGTGCTGAA TAAAACCCTA CTGCTCGCCG

                                       NotI
                                       ~~~~~~~~
  901  GCATGCACCT GGCCGGTCGC GGCCGCGGAA CCACTGAAGG ATGAGCTGTA
       CGTACGTGGA CCGGCCAGCG CCGGCGCCTT GGTGACTTCC TACTCGACAT

  951  AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT AAAGTTTCTT AAGATTGAAT
       TTCTTCGTCT AGCAAGTTTG TAAACCGTTA TTTCAAAGAA TTCTAACTTA

 1001  CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT TGAATTACGT
       GGACAACGGC CAGAACGCTA CTAATAGTAT ATTAAAGACA ACTTAATGCA

 1051  TAAGCATGTA ATAATTAACA TGTAATGCAT GACGTTATTT ATGAGATGGG
       ATTCGTACAT TATTAATTGT ACATTACGTA CTGCAATAAA TACTCTACCC

 1101  TTTTTATGAT TAGAGTCCCG CAATTATACA TTTAATACGC GATAGAAAAC
       AAAAATACTA ATCTCAGGGC GTTAATATGT AAATTATGCG CTATCTTTTG

 1151  AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG TGTCATCTAT
       TTTTATATCG CGCGTTTGAT CCTATTTAAT AGCGCGCGCC ACAGTAGATA

                      XbaI
                      ~~~~~~~
            ClaI  HindIII
            ~~~~~~~~~~~~~~
 1201  GTTACTAGAT CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC
       CAATGATCTA GCTATTCGAA GATCTCGCCG GCCACCTCGA GGTTAAGCGG

 1251  CTATAGTGAG TCGTATTACG CGCGCTCACT GGCCGTCGTT TTACAACGTC
       GATATCACTC AGCATAATGC GCGCGAGTGA CCGGCAGCAA AATGTTGCAG

 1301  GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT
       CACTGACCCT TTTGGGACCG CAATGGGTTG AATTAGCGGA ACGTCGTGTA

                      PvuII
                      ~~~~~~
 1351  CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC
```

Fig. 40 C

```
        GGGGGAAAGC GGTCGACCGC ATTATCGCTT CTCCGGGCGT GGCTAGCGGG

1401    TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGGACGCG CCCTGTAGCG
        AAGGGTTGTC AACGCGTCGG ACTTACCGCT TACCCTGCGC GGGACATCGC

1451    GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT GACCGCTACA
        CGCGTAATTC GCGCCGCCCA CACCACCAAT GCGCGTCGCA CTGGCGATGT

1501    CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT
        GAACGGTCGC GGGATCGCGG GCGAGGAAAG CGAAAGAAGG GAAGGAAAGA

1551    CGCCACGTTC GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT
        GCGGTGCAAG CGGCCGAAAG GGGCAGTTCG AGATTTAGCC CCCGAGGGAA

1601    TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT
        ATCCCAAGGC TAAATCACGA AATGCCGTGG AGCTGGGGTT TTTTGAACTA

1651    TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG
        ATCCCACTAC CAAGTGCATC ACCCGGTAGC GGGACTATCT GCCAAAAAGC

1701    CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA
        GGGAAACTGC AACCTCAGGT GCAAGAAATT ATCACCTGAG AACAAGGTTT

1751    CTGGAACAAC ACTCAACCCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG
        GACCTTGTTG TGAGTTGGGA TAGAGCCAGA TAAGAAAACT AAATATTCCC

1801    ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA
        TAAAACGGCT AAAGCCGGAT AACCAATTTT TTACTCGACT AAATTGTTTT

1851    ATTTAACGCG AATTTTAACA AAATATTAAC GCTTACAATT TAGGTGGCAC
        TAAATTGCGC TTAAAATTGT TTTATAATTG CGAATGTTAA ATCCACCGTG

1901    TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC
        AAAAGCCCCT TTACACGCGC CTTGGGGATA AACAAATAAA AAGATTTATG

1951    ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA
        TAAGTTTATA CATAGGCGAG TACTCTGTTA TTGGGACTAT TTACGAAGTT

2001    TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT
        ATTATAACTT TTTCCTTCTC ATACTCATAA GTTGTAAAGG CACAGCGGGA

2051    TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
        ATAAGGGAAA AAACGCCGTA AAACGGAAGG ACAAAAACGA GTGGGTCTTT

                                                ApaLI
                                                ~~~~~~~
2101    CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT
        GCGACCACTT TCATTTTCTA CGACTTCTAG TCAACCCACG TGCTCACCCA

2151    TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC
        ATGTAGCTTG ACCTAGAGTT GTCGCCATTC TAGGAACTCT CAAAAGCGGG

2201    CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG
        GCTTCTTGCA AAAGGTTACT ACTCGTGAAA ATTTCAAGAC GATACACCGC
```

Fig. 40D

2251 CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA
       GCCATAATAG GGCATAACTG CGGCCCGTTC TCGTTGAGCC AGCGGCGTAT

2301 CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA
       GTGATAAGAG TCTTACTGAA CCAACTCATG AGTGGTCAGT GTCTTTTCGT

2351 TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA
       AGAATGCCTA CCGTACTGTC ATTCTCTTAA TACGTCACGA CGGTATTGGT

2401 TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG
       ACTCACTATT GTGACGCCGG TTGAATGAAG ACTGTTGCTA GCCTCCTGGC

2451 AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT
       TTCCTCGATT GGCGAAAAAA CGTGTTGTAC CCCCTAGTAC ATTGAGCGGA

2501 TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG
       ACTAGCAACC CTTGGCCTCG ACTTACTTCG GTATGGTTTG CTGCTCGCAC

2551 ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT
       TGTGGTGCTA CGGACATCGT TACCGTTGTT GCAACGCGTT TGATAATTGA

2601 GGCGAACTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA
       CCGCTTGATG AATGAGATCG AAGGGCCGTT GTTAATTATC TGACCTACCT

2651 GGCGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT
       CCGCCTATTT CAACGTCCTG GTGAAGACGC GAGCCGGGAA GGCCGACCGA

2701 GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC
       CCAAATAACG ACTATTTAGA CCTCGGCCAC TCGCACCCAG AGCGCCATAG

2751 ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA
       TAACGTCGTG ACCCCGGTCT ACCATTCGGG AGGGCATAGC ATCAATAGAT

2801 CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG
       GTGCTGCCCC TCAGTCCGTT GATACCTACT TGCTTTATCT GTCTAGCGAC

2851 AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA CCAAGTTTAC
       TCTATCCACG GAGTGACTAA TTCGTAACCA TTGACAGTCT GGTTCAAATG

2901 TCATATATAC TTTAGATTGA TTTAAAACTT CATTTTTAAT TTAAAAGGAT
       AGTATATATG AAATCTAACT AAATTTTGAA GTAAAAATTA AATTTTCCTA

2951 CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG
       GATCCACTTC TAGGAAAAAC TATTAGAGTA CTGGTTTTAG GGAATTGCAC

3001 AGTTTTCGTT CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT
       TCAAAAGCAA GGTGACTCGC AGTCTGGGGC ATCTTTTCTA GTTTCCTAGA

3051 TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA
       AGAACTCTAG GAAAAAAAGA CGCGCATTAG ACGACGAACG TTTGTTTTTT

3101 ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC
       TGGTGGCGAT GGTCGCCACC AAACAAACGG CCTAGTTCTC GATGGTTGAG

Fig. 40E

```
3151  TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC
      AAAAAGGCTT CCATTGACCG AAGTCGTCTC GCGTCTATGG TTTATGACAG

3201  CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC
      GAAGATCACA TCGGCATCAA TCCGGTGGTG AAGTTCTTGA GACATCGTGG

3251  GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG
      CGGATGTATG GAGCGAGACG ATTAGGACAA TGGTCACCGA CGACGGTCAC

3301  GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT
      CGCTATTCAG CACAGAATGG CCCAACCTGA GTTCTGCTAT CAATGGCCTA

                                          ApaLI
                                          ~~~~~~
3351  AAGGCGCAGC GGTCGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT
      TTCCGCGTCG CCAGCCCGAC TTGCCCCCCA AGCACGTGTG TCGGGTCGAA

3401  GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG
      CCTCGCTTGC TGGATGTGGC TTGACTCTAT GGATGTCGCA CTCGATACTC

3451  AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC
      TTTCGCGGTG CGAAGGGCTT CCCTCTTTCC GCCTGTCCAT AGGCCATTCG

3501  GGCAGGGTCG AACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC
      CCGTCCCAGC CTTGTCCTCT CGCGTGCTCC CTCGAAGGTC CCCCTTTGCG

3551  CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC
      GACCATAGAA ATATCAGGAC AGCCCAAAGC GGTGGAGACT GAACTCGCAG

3601  GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC
      CTAAAAACAC TACGAGCAGT CCCCCCGCCT CGGATACCTT TTTGCGGTCG

3651  AACGCGGCCT TTTTACGGTT CCTGGCCTTT GCTGGCCTT TTGCTCACAT
      TTGCGCCGGA AAAATGCCAA GGACCGGAAA ACGACCGGAA AACGAGTGTA

3701  GTTCTTTCCT GCGTTATCCC CTGATTCTGT GGATAACCGT ATTACCGCCT
      CAAGAAAGGA CGCAATAGGG GACTAAGACA CCTATTGGCA TAATGGCGGA

3751  TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG
      AACTCACTCG ACTATGGCGA GCGGCGTCGG CTTGCTGGCT CGCGTCGCTC

3801  TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC CGCCTCTCCC
      AGTCACTCGC TCCTTCGCCT TCTCGCGGGT TATGCGTTTG GCGGAGAGGG

                                          PvuII
                                          ~~~~~~
3851  CGCGCGTTGG CCGATTCATT AATGCAGCTG GCACGACAGG TTTCCCGACT
      GCGCGCAACC GGCTAAGTAA TTACGTCGAC CGTGCTGTCC AAAGGGCTGA

3901  GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA GCTCACTCAT
      CCTTTCGCCC GTCACTCGCG TTGCGTTAAT TACACTCAAT CGAGTGAGTA

3951  TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTTGTGTGG
```

Fig. 40 F

```
           ATCCGTGGGG TCCGAAATGT GAAATACGAA GGCCGAGCAT ACAACACACC

     4001  AATTGTGAGC GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT
           TTAACACTCG CCTATTGTTA AAGTGTGTCC TTTGTCGATA CTGGTACTAA

     4051  ACGCCAAGCG CGCAATTAAC CCTCACTAAA GGGAACAAAA GCTGGGTACC
           TGCGGTTCGC GCGTTAATTG GGAGTGATTT CCCTTGTTTT CGACCCATGG

                      AvaI
                      ~~~~~~~
     4101  GGGCCCCCCC TCGAGGTCAT TCATATGCTT GAGAAGAGAG TCGGGATAGT
           CCCGGGGGGG AGCTCCAGTA AGTATACGAA CTCTTCTCTC AGCCCTATCA

     4151  CCAAAATAAA ACAAAGGTAA GATTACCTGG TCAAAAGTGA AAACATCAGT
           GGTTTTATTT TGTTTCCATT CTAATGGACC AGTTTTCACT TTTGTAGTCA

     4201  TAAAAGGTGG TATAAGTAAA ATATCGGTAA TAAAAGGTGG CCCAAAGTGA
           ATTTTCCACC ATATTCATTT TATAGCCATT ATTTTCCACC GGGTTTCACT

     4251  AATTTACTCT TTTCTACTAT TATAAAAATT GAGGATGTTT TGTCGGTACT
           TTAAATGAGA AAAGATGATA ATATTTTTAA CTCCTACAAA ACAGCCATGA

     4301  TTGATACGTC ATTTTTGTAT GAATTGGTTT TTAAGTTTAT TCGCGATTTG
           AACTATGCAG TAAAAACATA CTTAACCAAA AATTCAAATA AGCGCTAAAC

     4351  GAAATGCATA TCTGTATTTG AGTCGGTTTT TAAGTTCGTT GCTTTTGTAA
           CTTTACGTAT AGACATAAAC TCAGCCAAAA ATTCAAGCAA CGAAAACATT

     4401  ATACAGAGGG ATTTGTATAA GAAATATCTT TAAAAAACCC ATATGCTAAT
           TATGTCTCCC TAAACATATT CTTTATAGAA ATTTTTTGGG TATACGATTA

                                            EcoRI
                                            ~~~~~~~
     4451  TTGACATAAT TTTTGAGAAA AATATATATT CAGGCGAATT CCACAATGAA
           AACTGTATTA AAAACTCTTT TTATATATAA GTCCGCTTAA GGTGTTACTT

     4501  CAATAATAAG ATTAAAATAG CTTGCCCCCG TTGCAGCGAT GGGTATTTTT
           GTTATTATTC TAATTTTATC GAACGGGGGC AACGTCGCTA CCCATAAAAA

     4551  TCTAGTAAAA TAAAAGATAA ACTTAGACTC AAAACATTTA CAAAAACAAC
           AGATCATTTT ATTTTCTATT TGAATCTGAG TTTTGTAAAT GTTTTTGTTG

     4601  CCCTAAAGTC CTAAAGCCCA AAGTGCTATG CACGATCCAT AGCAAGCCCA
           GGGATTTCAG GATTTCGGGT TTCACGATAC GTGCTAGGTA TCGTTCGGGT

     4651  GCCCAACCCA ACCCAACCCA ACCCACCCCA GTGCAGCCAA CTGGCAAATA
           CGGGTTGGGT TGGGTTGGGT TGGGTGGGGT CACGTCGGTT GACCGTTTAT

     4701  GTCTCCACCC CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC
           CAGAGGTGGG GGCCGTGATA GTGGCACTCA ACAGGCGTGG TGGCGTGCAG

     4751  TCGCAGCCAA AAAAAAAAAA AGAAAGAAAA AAAAGAAAAA GAAAAACAGC
           AGCGTCGGTT TTTTTTTTTT TCTTTCTTTT TTTTCTTTTT CTTTTTGTCG
```

Fig. 40 G

```
4801  AGGTGGGTCC GGGTCGTGGG GGCCGGAAAA GCGAGGAGGA TCGCGAGCAG
      TCCACCCAGG CCCAGCACCC CCGGCCTTTT CGCTCCTCCT AGCGCTCGTC

4851  CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCCATCGCC
      GCTGCTCCGG GCCGGGAGGG AGGCGAAGGT TTCTTTGCGG GGGGTAGCGG

4901  ACTATATACA TACCCCCCCC TCTCCTCCCA TCCCCCCAAC CCTACCACCA
      TGATATATGT ATGGGGGGGG AGAGGAGGGT AGGGGGGTTG GGATGGTGGT

4951  CCACCACCAC CACCTCCTCC CCCCTCGCTG CCGGACGACG AGCTCCTCCC
      GGTGGTGGTG GTGGAGGAGG GGGGAGCGAC GGCCTGCTGC TCGAGGAGGG

5001  CCCTCCCCCT CCGCCGCCGC CGGTAACCAC CCCGCCCCTC TCCTCTTTCT
      GGGAGGGGGA GGCGGCGGCG GCCATTGGTG GGGCGGGGAG AGGAGAAAGA

5051  TTCTCCGTTT TTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT
      AAGAGGCAAA AAAAAAAGCA GAGCCAGAGC TAGAAACCGG AACCATCAAA

5101  GGGTGGGCGA GAGCGGCTTC GTCGCCCAGA TCGGTGCGCG GGAGGGGCGG
      CCCACCCGCT CTCGCCGAAG CAGCGGGTCT AGCCACGCGC CCTCCCCGCC

                                                  BamHI
                                                  ~~~~~~~
5151  GATCTCGCGG CTGGCGTCTC CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG
      CTAGAGCGCC GACCGCAGAG GCCCGCACTC AGCCGGGCCT AGGAGCGCCC

5201  GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG
      CTTACCCCGA GAGCCTACAT CTAGAAGAAA GAAAGAAGAA AAACACCATC

5251  AATTTGAATC CCTCAGCATT GTTCATCGGT AGTTTTTCTT TTCATGATTT
      TTAAACTTAG GGAGTCGTAA CAAGTAGCCA TCAAAAAGAA AAGTACTAAA

5301  GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGC
      CACTGTTTAC GTCGGAGCAC GCCTCGAAAA AACATCG
```

Fig. 40 H

Bgl II (4654)

Bam HI (4622)

Hin dIII (2)

Rice actin

Nco I (9)

sGFP

Not I (733)

Sma I (741)

Eco RI (3920)

Bss HII (3883)

Pst I (749)

Nos?

Bss HII (984)

pGT6
4773 bp

Bss HII (1008)

Cla I (1035)

Hin dIII (1040)

Xba I (1045)

Bss HII (1095)

ori

f1 ori

amp
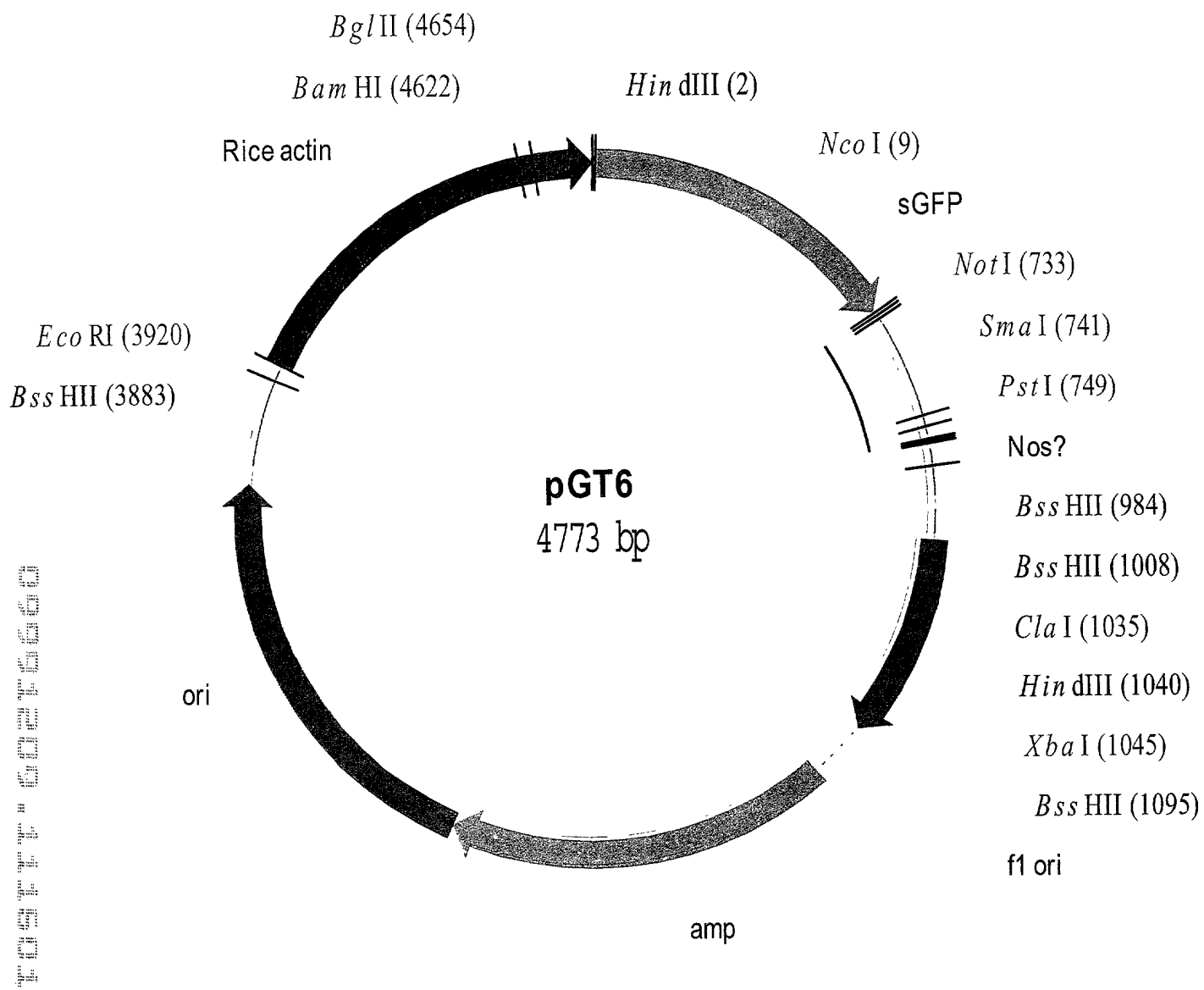
Fig. 41 A

Sequence for pGT6

```
         HindIII NcoI
         ~~~~~~ ~~~~~~~
     1   AAGCTTACCA TGGTGAGCAA GGGCGAGGAG CTGTTCACCG GGGTGGTGCC
CATCCTGGTC GAGCTGGACG
         TTCGAATGGT ACCACTCGTT CCCGCTCCTC GACAAGTGGC CCCACCACGG
GTAGGACCAG CTCGACCTGC

    71   GCGACGTGAA CGGCCACAAG TTCAGCGTGT CCGGCGAGGG CGAGGGCGAT
GCCACCTACG GCAAGCTGAC
         CGCTGCACTT GCCGGTGTTC AAGTCGCACA GGCCGCTCCC GCTCCCGCTA
CGGTGGATGC CGTTCGACTG

   141   CCTGAAGTTC ATCTGCACCA CCGGCAAGCT GCCCGTGCCC TGGCCCACCC
TCGTGACCAC CTTCACCTAC
         GGACTTCAAG TAGACGTGGT GGCCGTTCGA CGGGCACGGG ACCGGGTGGG
AGCACTGGTG GAAGTGGATG

   211   GGCGTGCAGT GCTTCAGCCG CTACCCCGAC CACATGAAGC AGCACGACTT
CTTCAAGTCC GCCATGCCCG
         CCGCACGTCA CGAAGTCGGC GATGGGGCTG GTGTACTTCG TCGTGCTGAA
GAAGTTCAGG CGGTACGGGC

   281   AAGGCTACGT CCAGGAGCGC ACCATCTTCT TCAAGGACGA CGGCAACTAC
AAGACCCGCG CCGAGGTGAA
         TTCCGATGCA GGTCCTCGCG TGGTAGAAGA AGTTCCTGCT GCCGTTGATG
TTCTGGGCGC GGCTCCACTT

   351   GTTCGAGGGC GACACCCTGG TGAACCGCAT CGAGCTGAAG GGCATCGACT
TCAAGGAGGA CGGCAACATC
         CAAGCTCCCG CTGTGGGACC ACTTGGCGTA GCTCGACTTC CCGTAGCTGA
AGTTCCTCCT GCCGTTGTAG

   421   CTGGGGCACA AGCTGGAGTA CAACTACAAC AGCCACAACG TCTATATCAT
GGCCGACAAG CAGAAGAACG
         GACCCCGTGT TCGACCTCAT GTTGATGTTG TCGGTGTTGC AGATATAGTA
CCGGCTGTTC GTCTTCTTGC

   491   GCATCAAGGT GAACTTCAAG ATCCGCCACA ACATCGAGGA CGGCAGCGTG
CAGCTCGCCG ACCACTACCA
         CGTAGTTCCA CTTGAAGTTC TAGGCGGTGT TGTAGCTCCT GCCGTCGCAC
GTCGAGCGGC TGGTGATGGT

   561   GCAGAACACC CCCATCGGCG ACGGCCCCGT GCTGCTGCCC GACAACCACT
ACCTGAGCAC CCAGTCCGCC
         CGTCTTGTGG GGGTAGCCGC TGCCGGGGCA CGACGACGGG CTGTTGGTGA
TGGACTCGTG GGTCAGGCGG

   631   CTGAGCAAAG ACCCCAACGA GAAGCGCGAT CACATGGTCC TGCTGGAGTT
CGTGACCGCC GCCGGGATCA
         GACTCGTTTC TGGGGTTGCT CTTCGCGCTA GTGTACCAGG ACGACCTCAA
GCACTGGCGG CGGCCCTAGT
```

Fig. 41B

```
                                              SmaI
                                            ~~~~~~~
                                    NotI            PstI
                                 ~~~~~~~~         ~~~~~~
      701   CTCACGGCAT GGACGAGCTG TACAAGTAAA GCGGCCGCCC GGGCTGCAGG
    GAAACCACTG AAGGATGAGC
            GAGTGCCGTA CCTGCTCGAC ATGTTCATTT CGCCGGCGGG CCCGACGTCC
    CTTTGGTGAC TTCCTACTCG

      771   TGTAAAGAAG CAGATCGTTC AAACATTTGG CAATAAAGTT TCTTAAGATT
    GAATCCTGTT GCCGGTCTTG
            ACATTTCTTC GTCTAGCAAG TTTGTAAACC GTTATTTCAA AGAATTCTAA
    CTTAGGACAA CGGCCAGAAC

      841   CGATGATTAT CATATAATTT CTGTTGAATT ACGTTAAGCA TGTAATAATT
    AACATGTAAT GCATGACGTT
            GCTACTAATA GTATATTAAA GACAACTTAA TGCAATTCGT ACATTATTAA
    TTGTACATTA CGTACTGCAA

      911   ATTTATGAGA TGGGTTTTTA TGATTAGAGT CCCGCAATTA TACATTTAAT
    ACGCGATAGA AAACAAAATA
            TAAATACTCT ACCCAAAAAT ACTAATCTCA GGGCGTTAAT ATGTAAATTA
    TGCGCTATCT TTTGTTTTAT


    XbaI              .

    ~~~~~~
              BssHII                    BssHII
    ClaI HindIII
             ~~~~~~                 ~~~~~~~
    ~~~~~~~~~~~~~~
             981   TAGCGCGCAA ACTAGGATAA ATTATCGCGC GCGGTGTCAT CTATGTTACT
    AGATCGATAA GCTTCTAGAG
            ATCGCGCGTT TGATCCTATT TAATAGCGCG CGCCACAGTA GATACAATGA
    TCTAGCTATT CGAAGATCTC

                                                        BssHII
                                                      ~~~~~~
     1051   CGGCCGGTGG AGCTCCAATT CGCCCTATAG TGAGTCGTAT TACGCGCGCT
    CACTGGCCGT CGTTTTACAA
            GCCGGCCACC TCGAGGTTAA GCGGGATATC ACTCAGCATA ATGCGCGCGA
    GTGACCGGCA GCAAAATGTT

     1121   CGTCGTGACT GGGAAAACCC TGGCGTTACC CAACTTAATC GCCTTGCAGC
    ACATCCCCCT TTCGCCAGCT
            GCAGCACTGA CCCTTTTGGG ACCGCAATGG GTTGAATTAG CGGAACGTCG
    TGTAGGGGGA AAGCGGTCGA

     1191   GGCGTAATAG CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCGC
    AGCCTGAATG GCGAATGGGA
            CCGCATTATC GCTTCTCCGG GCGTGGCTAG CGGGAAGGGT TGTCAACGCG
    TCGGACTTAC CGCTTACCCT
```

Fig. 41C

```
1261  CGCGCCCTGT AGCGGCGCAT TAAGCGCGGC GGGTGTGGTG GTTACGCGCA
GCGTGACCGC TACACTTGCC
          GCGCGGGACA TCGCCGCGTA ATTCGCGCCG CCCACACCAC CAATGCGCGT
CGCACTGGCG ATGTGAACGG

1331  AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC TTCCCTTCCT TTCTCGCCAC
GTTCGCCGGC TTTCCCCGTC
          TCGCGGGATC GCGGGCGAGG AAAGCGAAAG AAGGGAAGGA AAGAGCGGTG
CAAGCGGCCG AAAGGGGCAG

1401  AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT TCCGATTTAG TGCTTTACGG
CACCTCGACC CCAAAAAACT
          TTCGAGATTT AGCCCCCGAG GGAAATCCCA AGGCTAAATC ACGAAATGCC
GTGGAGCTGG GGTTTTTTGA

1471  TGATTAGGGT GATGGTTCAC GTAGTGGGCC ATCGCCCTGA TAGACGGTTT
TTCGCCCTTT GACGTTGGAG
          ACTAATCCCA CTACCAAGTG CATCACCCGG TAGCGGGACT ATCTGCCAAA
AAGCGGGAAA CTGCAACCTC

1541  TCCACGTTCT TTAATAGTGG ACTCTTGTTC CAAACTGGAA CAACACTCAA
CCCTATCTCG GTCTATTCTT
          AGGTGCAAGA AATTATCACC TGAGAACAAG GTTTGACCTT GTTGTGAGTT
GGGATAGAGC CAGATAAGAA

1611  TTGATTTATA AGGGATTTTG CCGATTTCGG CCTATTGGTT AAAAAATGAG
CTGATTTAAC AAAAATTTAA
          AACTAAATAT TCCCTAAAAC GGCTAAAGCC GGATAACCAA TTTTTTACTC
GACTAAATTG TTTTTAAATT

1681  CGCGAATTTT AACAAAATAT TAACGCTTAC AATTTAGGTG GCACTTTTCG
GGGAAATGTG CGCGGAACCC
          GCGCTTAAAA TTGTTTTATA ATTGCGAATG TTAAATCCAC CGTGAAAAGC
CCCTTTACAC GCGCCTTGGG

1751  CTATTTGTTT ATTTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA
CAATAACCCT GATAAATGCT
          GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAGG CGAGTACTCT
GTTATTGGGA CTATTTACGA

1821  TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG
CCCTTATTCC CTTTTTTGCG
          AGTTATTATA ACTTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC
GGGAATAAGG GAAAAAACGC

1891  GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA
AGATGCTGAA GATCAGTTGG
          CGTAAAACGG AAGGACAAAA ACGAGTGGGT CTTTGCGACC ACTTTCATTT
TCTACGACTT CTAGTCAACC

1961  GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT
GAGAGTTTTC GCCCCGAAGA
```

Fig. 41 D

```
        CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTCGCC ATTCTAGGAA
CTCTCAAAAG CGGGGCTTCT

   2031 ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT
TATCCCGTAT TGACGCCGGG
        TGCAAAAGGT TACTACTCGT GAAAATTTCA AGACGATACA CCGCGCCATA
ATAGGGCATA ACTGCGGCCC

   2101 CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG ACTTGGTTGA
GTACTCACCA GTCACAGAAA
        GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT
CATGAGTGGT CAGTGTCTTT

   2171 AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA
ACCATGAGTG ATAACACTGC
        TCGTAGAATG CCTACCGTAC TGTCATTCTC TTAATACGTC ACGACGGTAT
TGGTACTCAC TATTGTGACG

   2241 GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT
TTTTGCACAA CATGGGGGAT
        CCGGTTGAAT GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA
AAAACGTGTT GTACCCCCTA

   2311 CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC
AAACGACGAG CGTGACACCA
        GTACATTGAG CGGAACTAGC AACCCTTGGC CTCGACTTAC TTCGGTATGG
TTTGCTGCTC GCACTGTGGT

   2381 CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA
CTACTTACTC TAGCTTCCCG
        GCTACGGACA TCGTTACCGT TGTTGCAACG CGTTTGATAA TTGACCGCTT
GATGAATGAG ATCGAAGGGC

   2451 GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC
TGCGCTCGGC CCTTCCGGCT
        CGTTGTTAAT TATCTGACCT ACCTCCGCCT ATTTCAACGT CCTGGTGAAG
ACGCGAGCCG GGAAGGCCGA

   2521 GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG
TATCATTGCA GCACTGGGGC
        CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC
ATAGTAACGT CGTGACCCCG

   2591 CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG
GCAACTATGG ATGAACGAAA
        GTCTACCATT CGGGAGGGCA TAGCATCAAT AGATGTGCTG CCCCTCAGTC
CGTTGATACC TACTTGCTTT

   2661 TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT
CAGACCAAGT TTACTCATAT
        ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA
GTCTGGTTCA AATGAGTATA
```

Fig. 41 E

2731 ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT
GAAGATCCTT TTTGATAATC
        TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAAATTTT CCTAGATCCA
CTTCTAGGAA AAACTATTAG

2801 TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC
CCCGTAGAAA AGATCAAAGG
        AGTACTGGTT TTAGGGAATT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG
GGGCATCTTT TCTAGTTTCC

2871 ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA
AAAAACCACC GCTACCAGCG
        TAGAAGAACT CTAGGAAAAA AAGACGCGCA TTAGACGACG AACGTTTGTT
TTTTTGGTGG CGATGGTCGC

2941 GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC
TGGCTTCAGC AGAGCGCAGA
        CACCAAACAA ACGGCCTAGT TCTCGATGGT TGAGAAAAAG CTTCCATTG
ACCGAAGTCG TCTCGCGTCT

3011 TACCAAATAC TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG
AACTCTGTAG CACCGCCTAC
        ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC
TTGAGACATC GTGGCGGATG

3081 ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA
AGTCGTGTCT TACCGGGTTG
        TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT
TCAGCACAGA ATGGCCCAAC

3151 GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG
GGGTTCGTGC ACACAGCCCA
        CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC
CCCAAGCACG TGTGTCGGGT

3221 GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA
TGAGAAAGCG CCACGCTTCC
        CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGAT
ACTCTTTCGC GGTGCGAAGG

3291 CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG
GAGAGCGCAC GAGGGAGCTT
        GCTTCCCTCT TTCCGCCTGT CCATAGGCCA TTCGCCGTCC CAGCCTTGTC
CTCTCGCGTG CTCCCTCGAA

3361 CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT
CTGACTTGAG CGTCGATTTT
        GGTCCCCCTT TGCGGACCAT AGAAATATCA GGACAGCCCA AAGCGGTGGA
GACTGAACTC GCAGCTAAAA

3431 TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG
GCCTTTTTAC GGTTCCTGGC
        ACACTACGAG CAGTCCCCCC GCCTCGGATA CCTTTTTGCG GTCGTTGCGC
CGGAAAAATG CCAAGGACCG

Fig. 41 F

```
    3501  CTTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT
CTGTGGATAA CCGTATTACC
          GAAAACGACC GGAAAACGAG TGTACAAGAA AGGACGCAAT AGGGGACTAA
GACACCTATT GGCATAATGG

    3571  GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG
CGAGTCAGTG AGCGAGGAAG
          CGGAAACTCA CTCGACTATG GCGAGCGGCG TCGGCTTGCT GGCTCGCGTC
GCTCAGTCAC TCGCTCCTTC

    3641  CGGAAGAGCG CCCAATACGC AAACCGCCTC TCCCCGCGCG TTGGCCGATT
CATTAATGCA GCTGGCACGA
          GCCTTCTCGC GGGTTATGCG TTTGGCGGAG AGGGGCGCGC AACCGGCTAA
GTAATTACGT CGACCGTGCT

    3711  CAGGTTTCCC GACTGGAAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA
GTTAGCTCAC TCATTAGGCA
          GTCCAAAGGG CTGACCTTTC GCCCGTCACT CGCGTTGCGT TAATTACACT
CAATCGAGTG AGTAATCCGT

    3781  CCCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT GTGGAATTGT
GAGCGGATAA CAATTTCACA
          GGGGTCCGAA ATGTGAAATA CGAAGGCCGA GCATACAACA CACCTTAACA
CTCGCCTATT GTTAAAGTGT


                                           BssHII

EcoRI
                                           ~~~~~~

~~
    3851  CAGGAAACAG CTATGACCAT GATTACGCCA AGCGCGCAAT TAACCCTCAC
TAAAGGGAAC AAAAGCTGGA
          GTCCTTTGTC GATACTGGTA CTAATGCGGT TCGCGCGTTA ATTGGGAGTG
ATTTCCCTTG TTTTCGACCT

          EcoRI
          ~~~~
    3921  ATTCCACAAT GAACAATAAT AAGATTAAAA TAGCTTGCCC CCGTTGCAGC
GATGGGTATT TTTTCTAGTA
          TAAGGTGTTA CTTGTTATTA TTCTAATTTT ATCGAACGGG GGCAACGTCG
CTACCCATAA AAAAGATCAT

    3991  AAATAAAAGA TAAACTTAGA CTCAAACAT TTACAAAAAC AACCCCTAAA
GTCCTAAAGC CCAAAGTGCT
          TTTATTTTCT ATTTGAATCT GAGTTTTGTA AATGTTTTTG TTGGGGATTT
CAGGATTTCG GGTTTCACGA

    4061  ATGCACGATC CATAGCAAGC CCAGCCCAAC CCAACCCAAC CCAACCCACC
CCAGTGCAGC CAACTGGCAA
          TACGTGCTAG GTATCGTTCG GGTCGGGTTG GGTTGGGTTG GGTTGGGTGG
GGTCACGTCG GTTGACCGTT

    4131  ATAGTCTCCA CCCCCGGCAC TATCACCGTG AGTTGTCCGC ACCACCGCAC
GTCTCGCAGC CAAAAAAAAA
```

Fig. 41 G

TATCAGAGGT GGGGGCCGTG ATAGTGGCAC TCAACAGGCG TGGTGGCGTG
CAGAGCGTCG GTTTTTTTTT

4201 AAAAGAAAGA AAAAAAGAA AAAGAAAAAC AGCAGGTGGG TCCGGGTCGT
GGGGGCCGGA AAAGCGAGGA
        TTTTCTTTCT TTTTTTTCTT TTTCTTTTTG TCGTCCACCC AGGCCCAGCA
CCCCCGGCCT TTTCGCTCCT

4271 GGATCGCGAG CAGCGACGAG GCCCGGCCCT CCCTCCGCTT CCAAAGAAAC
GCCCCCCATC GCCACTATAT
        CCTAGCGCTC GTCGCTGCTC CGGGCCGGGA GGGAGGCGAA GGTTTCTTTG
CGGGGGGTAG CGGTGATATA

4341 ACATACCCCC CCCTCTCCTC CCATCCCCCC AACCCTACCA CCACCACCAC
CACCACCTCC TCCCCCCTCG
        TGTATGGGGG GGGAGAGGAG GGTAGGGGGG TTGGGATGGT GGTGGTGGTG
GTGGTGGAGG AGGGGGGAGC

4411 CTGCCGGACG ACGAGCTCCT CCCCCCTCCC CCTCCGCCGC CGCCGGTAAC
CACCCCGCCC CTCTCCTCTT
        GACGGCCTGC TGCTCGAGGA GGGGGGAGGG GGAGGCGGCG GCGGCCATTG
GTGGGGCGGG GAGAGGAGAA

4481 TCTTTCTCCG TTTTTTTTTT CGTCTCGGTC TCGATCTTTG GCCTTGGTAG
TTTGGGTGGG CGAGAGCGGC
        AGAAAGAGGC AAAAAAAAAA GCAGAGCCAG AGCTAGAAAC CGGAACCATC
AAACCCACCC GCTCTCGCCG

4551 TTCGTCGCCC AGATCGGTGC GCGGGAGGGG CGGGATCTCG CGGCTGGCGT
CTCCGGGCGT GAGTCGGCCC
        AAGCAGCGGG TCTAGCCACG CGCCCTCCCC GCCCTAGAGC GCCGACCGCA
GAGGCCCGCA CTCAGCCGGG

        BamHI   BglII
        ~~~~~~   ~~~~~~
4621 GGATCCTCGC GGGGAATGGG GCTCTCGGAT GTAGATCTTC TTTCTTTCTT
CTTTTTGTGG TAGAATTTGA
        CCTAGGAGCG CCCCTTACCC CGAGAGCCTA CATCTAGAAG AAAGAAAGAA
GAAAAACACC ATCTTAAACT

4691 ATCCCTCAGC ATTGTTCATC GGTAGTTTTT CTTTTCATGA TTTGTGACAA
ATGCAGCCTC GTGCGGAGCT
        TAGGGAGTCG TAACAAGTAG CCATCAAAAA GAAAGTACT AAACACTGTT
TACGTCGGAG CACGCCTCGA

4761 TTTTTGTAGG TAG
        AAAAACATCC ATC

Fig. 41 H

Kpn I (1)

Nco I (1)

Bss HII (4994)   Pst I (91)

Xba I (113)

Xba I (121)

Sma I (367)

ori   SEE promoter

Sph I (523)

Pst I (724)

pJQ5   PPI
5034 bp
Nco I (945)

aleurain signal/NPIR

amp   Not I (1014)

Sal I (1114)

Acc I (1115)

Bam HI (1170)

f1 ori   Glycosylation

Intron 1

FAE

Xho I (1552)

Nco I (1639)

Redundancy (stop)

KDEL

nos

Bss HII (2095)

Bss HII (2119)

Cla I (2146)

Hin dIII (2151)

Xba I (2156)

Bss HII (2206)

Fig. 42A

Sequence for pJQ5

```
      NcoI
      ~~~~~
      KpnI
      ~
   1  CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC
ACCATTGGCT ACAATATCTG
      GTACCCGGTC CATATTAATA CCCTATAGAG TTCGTTTATT AGCTTTATAG
TGGTAACCGA TGTTATAGAC


                     PstI                          XbaI      XbaI
                     ~~~~~~~                        ~~~~~~    ~~~~~~~
  71  AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTTGT ATCTAGAACT
CTAGATAGCA CAGCCACAGC
      TCGAGGCTCA AGACTGACGT CAGACCTACT GCGCACAACA TAGATCTTGA
GATCTATCGT GTCGGTGTCG


 141  ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTTG GAGACGGAGC
TCTTTCCTAC CTCCTGACGT
      TGGATGTCCT CACGCTGTGA ACACCTGACA TCATCACAAC CTCTGCCTCG
AGAAAGGATG GAGGACTGCA


 211  TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC
CCAACAAAAT ATCGTCCCCC
      ACGGCGGCAA CAGGTAAGGT TGCCGTAGTG AGAGTTGGTT AGTGCGCGAG
GGTTGTTTTA TAGCAGGGGG


 281  ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTCGCGCC GTTTTTGTCT
GAATCTCGCT TCCACTGGCC
      TACAGAACCG CCTCTCTCTC ATGTATGTAC GACAGCGCGG CAAAAACAGA
CTTAGAGCGA AGGTGACCGG


                          SmaI
                          ~~~~~~
 351  AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCG
TCACCCCTGG CGTCATGGGA
      TTAGTCGAGT CGAGGGCCCT CGAGTGAGTA AGTTCTAGGG TAGCAGCAGC
AGTGGGGACC GCAGTACCCT
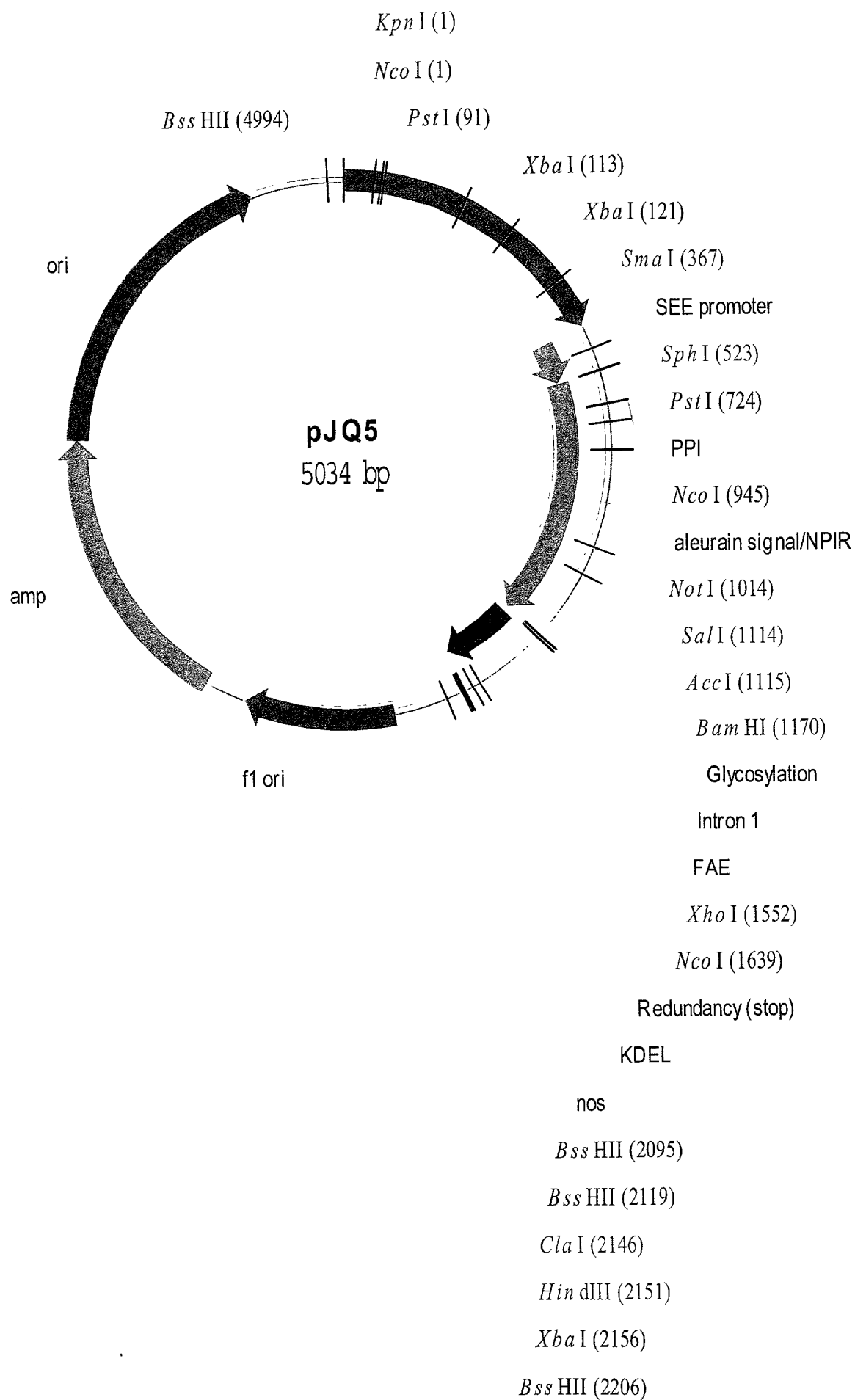

 421  TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA
GTACTGCAAG ATAACCCAAT
      ACCTTTTCTT GGAGGCAACG AGCCTACTCA GTCGGTATAG GGGCTTGTCT
CATGACGTTC TATTGGGTTA


                                     SphI
                                     ~~~~~~~
 491  TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTTTGG
CTTAATTGAC TTTATTTTTG
      AGTCTAAGGG GGTTATCTCT TTCATATCGT ACGAAAGCCC AAAACAAACC
GAATTAACTG AAATAAAAAC
```

Fig. 42B

561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT
CGAGACGGAT AATAGGCTGG
         AACCTCAACT TACGACTAAA CAACACATTT TACGGGTTGG TAGACTTATA
GCTCTGCCTA TTATCCGACC

631 CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT
GGGCATTACA GCTGGAGGCT
         GATTAATTAA ATATCGTTCT AAGACATCAC GTGTAGCGTT TATAGAAAGA
CCCGTAATGT CGACCTCCGA

                              PstI
                            ~~~~~~~
701 TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG
CGATGAGATG GGTATAAAAC
         AGTAGTCGGA CTTTGTGAGA CGTCTCGGAC TTCGTTCACC ACTTCGCACC
GCTACTCTAC CCATATTTTG

771 CCCCGGCACC GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCTCGC
TCCCCCTGCC GGACGACCCA
         GGGGCCGTGG CCCTGCGCTC GAGGGCGGAT GGTCATGGTA GAGCGGAGCG
AGGGGGACGG CCTGCTGGGT

841 GTAAAATACT GTTGCCCACT CGCCGGCGAG ATGGMCGTGC ACAAGGAGGT
SAACTTCGTS GCCTACCTCC
         CATTTTATGA CAACGGGTGA GCGGCCGCTC TACCKGCACG TGTTCCTCCA
STTGAAGCAS CGGATGGAGG

                                       NcoI
                                     ~~~~~~
911 TGATCGTSCT CGGCCTCCTC TTGCTCGTST CCGCCATGGA GCACGTGGAC
GCCAAGGCCT GCACCCKCGA
         ACTAGCASGA GCCGGAGGAG AACGAGCASA GGCGGTACCT CGTGCACCTG
CGGTTCCGGA CGTGGGMGCT

                                       NotI
                                     ~~~~~~~~
981 GTGCGGCAAC CTCGGCTTCG GCATCTGCCC GGCGGCCGCC TCCACGCAGG
GCATCTCCGA AGACCTCTAC
         CACGCCGTTG GAGCCGAAGC CGTAGACGGG CCGCCGGCGG AGGTGCGTCC
CGTAGAGGCT TCTGGAGATG

SalI

~~~~~~

AccI

~~~~~~
1051 AGCCGTTTAG TCGAAATGGC CACTATCTCC CAAGCTGCCT ACGCCGACCT
GTGCAACATT CCGTCGACTA
         TCGGCAAATC AGCTTTACCG GTGATAGAGG GTTCGACGGA TGCGGCTGGA
CACGTTGTAA GGCAGCTGAT

Fig. 42 C

BamHI
~~~~~~~

1121 TTATCAAGGG AGAGAAAATT TACAATTCTC AAACTGACAT TAACGGATGG
ATCCTCCGCG ACGACAGCAG
          AATAGTTCCC TCTCTTTTAA ATGTTAAGAG TTTGACTGTA ATTGCCTACC
TAGGAGGCGC TGCTGTCGTC

1191 CAAAGAAATA ATCACCGTCT TCCGTGGCAC TGGTAGTGAT ACGAATCTAC
AACTCGATAC TAACTACACC
          GTTTCTTTAT TAGTGGCAGA AGGCACCGTG ACCATCACTA TGCTTAGATG
TTGAGCTATG ATTGATGTGG

1261 CTCACGCCTT TCGACACCCT ACCACAATGC AACGGTTGTG AAGTACACGG
TGGATATTAT ATTGGATGGG
          GAGTGCGGAA AGCTGTGGGA TGGTGTTACG TTGCCAACAC TTCATGTGCC
ACCTATAATA TAACCTACCC

1331 TCTCCGTCCA GGACCAAGTC GAGTCGCTTG TCAAACAGCA GGTTAGCCAG
TATCCGGACT ACGCGCTGAC
          AGAGGCAGGT CCTGGTTCAG CTCAGCGAAC AGTTTGTCGT CCAATCGGTC
ATAGGCCTGA TGCGCGACTG

1401 CGTGACCGGC CACKCCCTCG GCGCCTCCCT GGCGGCACTC ACTGCCGCCC
AGCTGTCTGC GACATACGAC
          GCACTGGCCG GTGMGGGAGC CGCGGAGGGA CCGCCGTGAG TGACGGCGGG
TCGACAGACG CTGTATGCTG

1471 AACATCCGCC TGTACACCTT CGGCGAACCG CGCAGCGGCA ATCAGGCCTT
CGCGTCGTAC ATGAACGATG
          TTGTAGGCGG ACATGTGGAA GCCGCTTGGC GCGTCGCCGT TAGTCCGGAA
GCGCAGCATG TACTTGCTAC

XhoI
~~~~~~

1541 CCTTCCAAGC CTCGAGCCCA GATACGACGC AGTATTTCCG GGTCACTCAT
GCCAACGACG GCATCCCAAA
          GGAAGGTTCG GAGCTCGGGT CTATGCTGCG TCATAAAGGC CCAGTGAGTA
CGGTTGCTGC CGTAGGGTTT

NcoI
~~~~~~~

1611 CCTGCCCCCG GTGGAGCAGG GGTACGCCCA TGGCGGTGTA GAGTACTGGA
GCGTTGATCC TTACAGCGCC
          GGACGGGGGC CACCTCGTCC CCATGCGGGT ACCGCCACAT CTCATGACCT
CGCAACTAGG AATGTCGCGG

1681 CAGAACACAT TTGTCTGCAC TGGGGATGAA GTGCAGTGCT GTGAGGCCCA
GGGCGGACAG GGTGTGAATA
          GTCTTGTGTA AACAGACGTG ACCCCTACTT CACGTCACGA CACTCCGGGT
CCCGCCTGTC CCACACTTAT

1751 ATGCGCACAC GACTTATTTT GGGATGACGA GCGGAGCCTG TACATGGTGA
TCAGTCATTT CAGCCTCCCC

Fig. 42 D

```
         TACGCGTGTG CTGAATAAAA CCCTACTGCT CGCCTCGGAC ATGTACCACT
AGTCAGTAAA GTCGGAGGGG

   1821  GAGTGTACCA GGAAAGATGG ATGTCCTGGA GAGGGGGCCG CGTAACCACT
GAAGGATGAG CTGTAAAGAA
         CTCACATGGT CCTTTCTACC TACAGGACCT CTCCCCCGGC GCATTGGTGA
CTTCCTACTC GACATTTCTT

   1891  GCAGATCGTT CAAACATTTG GCAATAAAGT TTCTTAAGAT TGAATCCTGT
TGCCGGTCTT GCGATGATTA
         CGTCTAGCAA GTTTGTAAAC CGTTATTTCA AGAATTCTA ACTTAGGACA
ACGGCCAGAA CGCTACTAAT

   1961  TCATATAATT TCTGTTGAAT TACGTTAAGC ATGTAATAAT TAACATGTAA
TGCATGACGT TATTTATGAG
         AGTATATTAA AGACAACTTA ATGCAATTCG TACATTATTA ATTGTACATT
ACGTACTGCA ATAAATACTC


BssHII

~~~~~~
   2031  ATGGGTTTTT ATGATTAGAG TCCCGCAATT ATACATTTAA TACGCGATAG
AAAACAAAAT ATAGCGCGCA
         TACCCAAAAA TACTAATCTC AGGGCGTTAA TATGTAAATT ATGCGCTATC
TTTTGTTTTA TATCGCGCGT


XbaI

~~~~~~
                              BssHII                       ClaI HindIII
                              ~~~~~~                       ~~~~~~~~~~~~~
   2101  AACTAGGATA AATTATCGCG CGCGGTGTCA TCTATGTTAC TAGATCGATA
AGCTTCTAGA GCGGCCGGTG
         TTGATCCTAT TTAATAGCGC GCGCCACAGT AGATACAATG ATCTAGCTAT
TCGAAGATCT CGCCGGCCAC


                                                  BssHII
                                                  ~~~~~~
   2171  GAGCTCCAAT TCGCCCTATA GTGAGTCGTA TTACGCGCGC TCACTGGCCG
TCGTTTTACA ACGTCGTGAC
         CTCGAGGTTA AGCGGGATAT CACTCAGCAT AATGCGCGCG AGTGACCGGC
AGCAAAATGT TGCAGCACTG

   2241  TGGGAAAACC CTGGCGTTAC CCAACTTAAT CGCCTTGCAG CACATCCCCC
TTTCGCCAGC TGGCGTAATA
         ACCCTTTTGG GACCGCAATG GGTTGAATTA GCGGAACGTC GTGTAGGGGG
AAAGCGGTCG ACCGCATTAT

   2311  GCGAAGAGGC CCGCACCGAT CGCCCTTCCC AACAGTTGCG CAGCCTGAAT
GGCGAATGGG ACGCGCCCTG
         CGCTTCTCCG GGCGTGGCTA GCGGGAAGGG TTGTCAACGC GTCGGACTTA
CCGCTTACCC TGCGCGGGAC
```

Fig. 42 E

2381 TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT GGTTACGCGC AGCGTGACCG
CTACACTTGC CAGCGCCCTA
ATCGCCGCGT AATTCGCGCC GCCCACACCA CCAATGCGCG TCGCACTGGC
GATGTGAACG GTCGCGGGAT

2451 GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC TTTCTCGCCA CGTTCGCCGG
CTTTCCCCGT CAAGCTCTAA
CGCGGGCGAG GAAAGCGAAA GAAGGGAAGG AAAGAGCGGT GCAAGCGGCC
GAAAGGGGCA GTTCGAGATT

2521 ATCGGGGGCT CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC
CCCAAAAAAC TTGATTAGGG
TAGCCCCCGA GGGAAATCCC AAGGCTAAAT CACGAAATGC CGTGGAGCTG
GGGTTTTTTG AACTAATCCC

2591 TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTCGCCCTT
TGACGTTGGA GTCCACGTTC
ACTACCAAGT GCATCACCCG GTAGCGGGAC TATCTGCCAA AAAGCGGGAA
ACTGCAACCT CAGGTGCAAG

2661 TTTAATAGTG GACTCTTGTT CCAAACTGGA ACAACACTCA ACCCTATCTC
GGTCTATTCT TTTGATTTAT
AAATTATCAC CTGAGAACAA GGTTTGACCT TGTTGTGAGT TGGGATAGAG
CCAGATAAGA AAACTAAATA

2731 AAGGGATTTT GCCGATTTCG GCCTATTGGT TAAAAAATGA GCTGATTTAA
CAAAAATTTA ACGCGAATTT
TTCCCTAAAA CGGCTAAAGC CGGATAACCA ATTTTTTACT CGACTAAATT
GTTTTTAAAT TGCGCTTAAA

2801 TAACAAAATA TTAACGCTTA CAATTTAGGT GGCACTTTTC GGGGAAATGT
GCGCGGAACC CCTATTTGTT
ATTGTTTTAT AATTGCGAAT GTTAAATCCA CCGTGAAAAG CCCCTTTACA
CGCGCCTTGG GGATAAACAA

2871 TATTTTTCTA AATACATTCA AATATGTATC CGCTCATGAG ACAATAACCC
TGATAAATGC TTCAATAATA
ATAAAAAGAT TTATGTAAGT TTATACATAG GCGAGTACTC TGTTATTGGG
ACTATTTACG AAGTTATTAT

2941 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCGTGTC GCCCTTATTC
CCTTTTTTGC GGCATTTTGC
AACTTTTTCC TTCTCATACT CATAAGTTGT AAAGGCACAG CGGGAATAAG
GGAAAAAACG CCGTAAAACG

3011 CTTCCTGTTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA
AGATCAGTTG GGTGCACGAG
GAAGGACAAA AACGAGTGGG TCTTTGCGAC CACTTTCATT TTCTACGACT
TCTAGTCAAC CCACGTGCTC

3081 TGGGTTACAT CGAACTGGAT CTCAACAGCG GTAAGATCCT TGAGAGTTTT
CGCCCCGAAG AACGTTTTCC

Fig. 42 F

ACCCAATGTA GCTTGACCTA GAGTTGTCGC CATTCTAGGA ACTCTCAAAA
GCGGGGCTTC TTGCAAAAGG

    3151   AATGATGAGC ACTTTTAAAG TTCTGCTATG TGGCGCGGTA TTATCCCGTA
TTGACGCCGG GCAAGAGCAA
           TTACTACTCG TGAAAATTTC AAGACGATAC ACCGCGCCAT AATAGGGCAT
AACTGCGGCC CGTTCTCGTT

    3221   CTCGGTCGCC GCATACACTA TTCTCAGAAT GACTTGGTTG AGTACTCACC
AGTCACAGAA AAGCATCTTA
           GAGCCAGCGG CGTATGTGAT AAGAGTCTTA CTGAACCAAC TCATGAGTGG
TCAGTGTCTT TTCGTAGAAT

    3291   CGGATGGCAT GACAGTAAGA GAATTATGCA GTGCTGCCAT AACCATGAGT
GATAACACTG CGGCCAACTT
           GCCTACCGTA CTGTCATTCT CTTAATACGT CACGACGGTA TTGGTACTCA
CTATTGTGAC GCCGGTTGAA

    3361   ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTTGCACA
ACATGGGGGA TCATGTAACT
           TGAAGACTGT TGCTAGCCTC CTGGCTTCCT CGATTGGCGA AAAAACGTGT
TGTACCCCCT AGTACATTGA

    3431   CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA
GCGTGACACC ACGATGCCTG
           GCGGAACTAG CAACCCTTGG CCTCGACTTA CTTCGGTATG GTTTGCTGCT
CGCACTGTGG TGCTACGGAC

    3501   TAGCAATGGC AACAACGTTG CGCAAACTAT TAACTGGCGA ACTACTTACT
CTAGCTTCCC GGCAACAATT
           ATCGTTACCG TTGTTGCAAC GCGTTTGATA ATTGACCGCT TGATGAATGA
GATCGAAGGG CCGTTGTTAA

    3571   AATAGACTGG ATGGAGGCGG ATAAAGTTGC AGGACCACTT CTGCGCTCGG
CCCTTCCGGC TGGCTGGTTT
           TTATCTGACC TACCTCCGCC TATTTCAACG TCCTGGTGAA GACGCGAGCC
GGGAAGGCCG ACCGACCAAA

    3641   ATTGCTGATA AATCTGGAGC CGGTGAGCGT GGGTCTCGCG GTATCATTGC
AGCACTGGGG CCAGATGGTA
           TAACGACTAT TTAGACCTCG GCCACTCGCA CCCAGAGCGC CATAGTAACG
TCGTGACCCC GGTCTACCAT

    3711   AGCCCTCCCG TATCGTAGTT ATCTACACGA CGGGGAGTCA GGCAACTATG
GATGAACGAA ATAGACAGAT
           TCGGGAGGGC ATAGCATCAA TAGATGTGCT GCCCCTCAGT CCGTTGATAC
CTACTTGCTT TATCTGTCTA

    3781   CGCTGAGATA GGTGCCTCAC TGATTAAGCA TTGGTAACTG TCAGACCAAG
TTTACTCATA TATACTTTAG
           GCGACTCTAT CCACGGAGTG ACTAATTCGT AACCATTGAC AGTCTGGTTC
AAATGAGTAT ATATGAAATC

                            Fig. 42 G

3851   ATTGATTTAA AACTTCATTT TTAATTTAAA AGGATCTAGG TGAAGATCCT
TTTTGATAAT CTCATGACCA
           TAACTAAATT TTGAAGTAAA AATTAAATTT TCCTAGATCC ACTTCTAGGA
AAAACTATTA GAGTACTGGT

3921   AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGCGTCAGA CCCCGTAGAA
AAGATCAAAG GATCTTCTTG
           TTTAGGGAAT TGCACTCAAA AGCAAGGTGA CTCGCAGTCT GGGGCATCTT
TTCTAGTTTC CTAGAAGAAC

3991   AGATCCTTTT TTTCTGCGCG TAATCTGCTG CTTGCAAACA AAAAAACCAC
CGCTACCAGC GGTGGTTTGT
           TCTAGGAAAA AAAGACGCGC ATTAGACGAC GAACGTTTGT TTTTTTGGTG
GCGATGGTCG CCACCAAACA

4061   TTGCCGGATC AAGAGCTACC AACTCTTTTT CCGAAGGTAA CTGGCTTCAG
CAGAGCGCAG ATACCAAATA
           AACGGCCTAG TTCTCGATGG TTGAGAAAAA GGCTTCCATT GACCGAAGTC
GTCTCGCGTC TATGGTTTAT

4131   CTGTCCTTCT AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA
GCACCGCCTA CATACCTCGC
           GACAGGAAGA TCACATCGGC ATCAATCCGG TGGTGAAGTT CTTGAGACAT
CGTGGCGGAT GTATGGAGCG

4201   TCTGCTAATC CTGTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC
TTACCGGGTT GGACTCAAGA
           AGACGATTAG GACAATGGTC ACCGACGACG GTCACCGCTA TTCAGCACAG
AATGGCCCAA CCTGAGTTCT

4271   CGATAGTTAC CGGATAAGGC GCAGCGGTCG GGCTGAACGG GGGGTTCGTG
CACACAGCCC AGCTTGGAGC
           GCTATCAATG GCCTATTCCG CGTCGCCAGC CCGACTTGCC CCCCAAGCAC
GTGTGTCGGG TCGAACCTCG

4341   GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCT ATGAGAAAGC
GCCACGCTTC CCGAAGGGAG
           CTTGCTGGAT GTGGCTTGAC TCTATGGATG TCGCACTCGA TACTCTTTCG
CGGTGCGAAG GGCTTCCCTC

4411   AAAGGCGGAC AGGTATCCGG TAAGCGGCAG GGTCGGAACA GGAGAGCGCA
CGAGGGAGCT TCCAGGGGGA
           TTTCCGCCTG TCCATAGGCC ATTCGCCGTC CCAGCCTTGT CCTCTCGCGT
GCTCCCTCGA AGGTCCCCCT

4481   AACGCCTGGT ATCTTTATAG TCCTGTCGGG TTTCGCCACC TCTGACTTGA
GCGTCGATTT TTGTGATGCT
           TTGCGGACCA TAGAAATATC AGGACAGCCC AAAGCGGTGG AGACTGAACT
CGCAGCTAAA AACACTACGA

4551   CGTCAGGGGG GCGGAGCCTA TGGAAAAACG CCAGCAACGC GGCCTTTTTA
CGGTTCCTGG CCTTTTGCTG
           GCAGTCCCCC CGCCTCGGAT ACCTTTTTGC GGTCGTTGCG CCGGAAAAAT
GCCAAGGACC GGAAAACGAC

Fig. 42 H

4621   GCCTTTTGCT CACATGTTCT TTCCTGCGTT ATCCCCTGAT TCTGTGGATA
ACCGTATTAC CGCCTTTGAG
        CGGAAAACGA GTGTACAAGA AAGGACGCAA TAGGGGACTA AGACACCTAT
TGGCATAATG GCGGAAACTC

4691   TGAGCTGATA CCGCTCGCCG CAGCCGAACG ACCGAGCGCA GCGAGTCAGT
GAGCGAGGAA GCGGAAGAGC
        ACTCGACTAT GGCGAGCGGC GTCGGCTTGC TGGCTCGCGT CGCTCAGTCA
CTCGCTCCTT CGCCTTCTCG

4761   GCCCAATACG CAAACCGCCT CTCCCCGCGC GTTGGCCGAT TCATTAATGC
AGCTGGCACG ACAGGTTTCC
        CGGGTTATGC GTTTGGCGGA GAGGGGCGCG CAACCGGCTA AGTAATTACG
TCGACCGTGC TGTCCAAAGG

4831   CGACTGGAAA GCGGGCAGTG AGCGCAACGC AATTAATGTG AGTTAGCTCA
CTCATTAGGC ACCCCAGGCT
        GCTGACCTTT CGCCCGTCAC TCGCGTTGCG TTAATTACAC TCAATCGAGT
GAGTAATCCG TGGGGTCCGA

4901   TTACACTTTA TGCTTCCGGC TCGTATGTTG TGTGGAATTG TGAGCGGATA
ACAATTTCAC ACAGGAAACA
        AATGTGAAAT ACGAAGGCCG AGCATACAAC ACACCTTAAC ACTCGCCTAT
TGTTAAAGTG TGTCCTTTGT

NcoI

                                    BssHII

KpnI
                                    ~~~~~~
        ~~~~~
    4971    GCTATGACCA TGATTACGCC AAGCGCGCAA TTAACCCTCA CTAAAGGGAA
CAAAAGCTGG GTAC
        CGATACTGGT ACTAATGCGG TTCGCGCGTT AATTGGGAGT GATTTCCCTT
GTTTTCGACC CATG

Fig. 42 I

*Hin* dIII (2)

*Bgl* II (4831)      *Nco* I (9)

*Bam* HI (4799)      *Not* I (129)

Rice actin                        *Sal* I (229)

*Acc* I (230)

*Eco* RI (4097)                        *Bam* HI (285)

*Bss* HII (4060)                        SIG-VAC-FAE-F/S

*Xho* I (667)

**pJO6.1**

**4950 bp**                 *Nco* I (754)

*Sph* I (906)

*Not* I (921)

*Bss* HII (1161)

*Bss* HII (1185)

ori                                  *Cla* I (1212)

*Hin* dIII (1217)

*Xba* I (1222)

*Bss* HII (1272)

f1 ori

amp

Fig. 43A

# Sequence for pJO6

```
          HindIII NcoI
          ~~~~~~ ~~~~~~~
    1  AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT
GGCCACGGCC GCCGTCGCCG
          TTCGAATGGT ACCGGGTGCG GGCGCAGGAG GAGGACCGCG AGCGGCACGA
CCGGTGCCGG CGGCAGCGGC


NotI

~~~~~~~~~
    71  TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC
GACCGCGCGG CCGCCTCCAC
          AGCGGAGGAG GAGGAGGAAG CGGCTGAGGT TGGGCTAGGC CGGGCAGTGG
CTGGCGCGCC GGCGGAGGTG
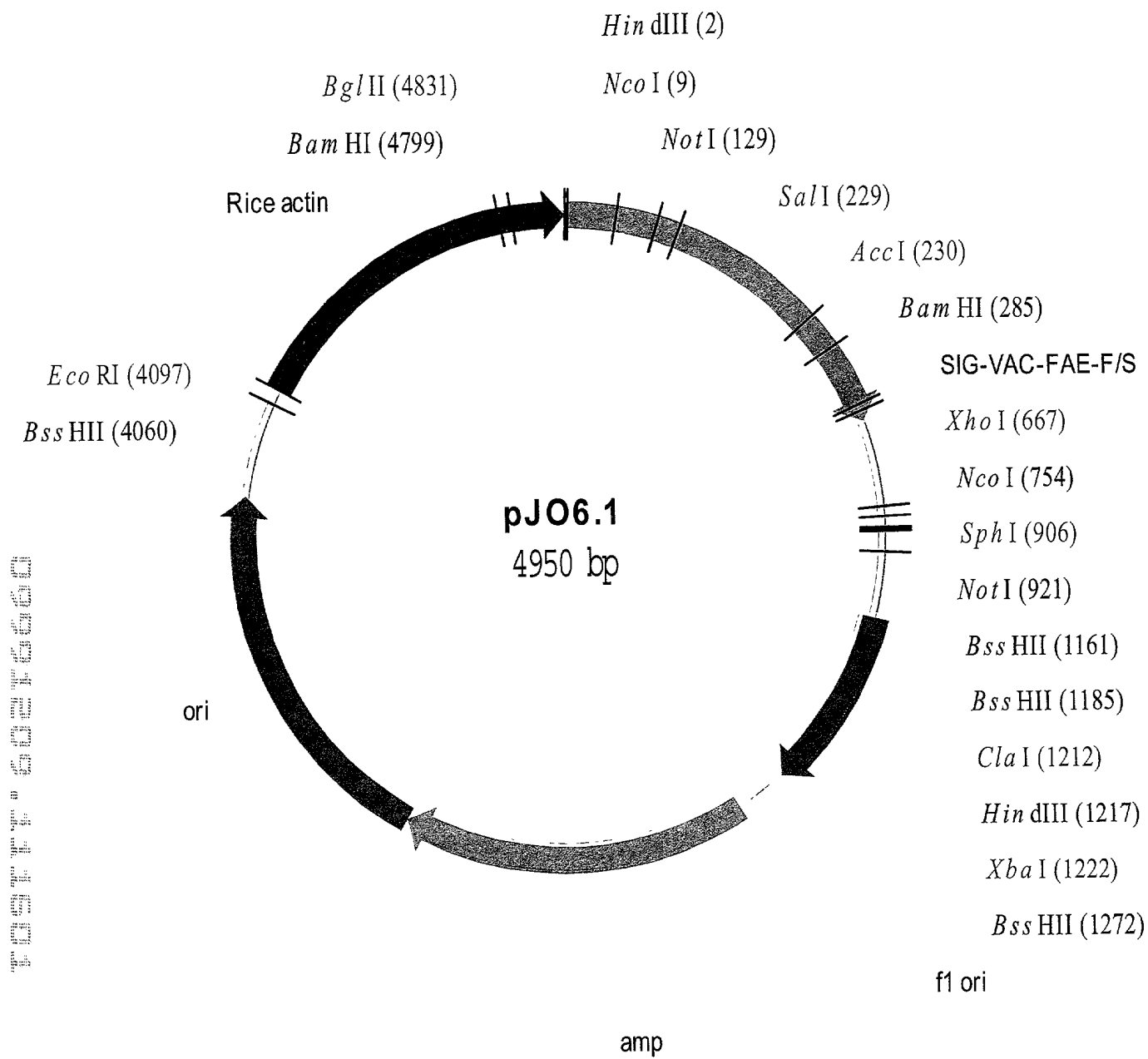
   141  GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA
TCTCCCAAGC TGCCTACGCC
          CGTCCCGTAG AGGCTTCTGG AGATGTCGGC AAATCAGCTT TACCGGTGAT
AGAGGGTTCG ACGGATGCGG


                              SalI
                              ~~~~~~~
                              AccI
                              ~~~~~~~
   211  GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTTACAA
TTCTCAAACT GACATTAACG
          CTGGACACGT TGTAAGGCAG CTGATAATAG TTCCCTCTCT TTTAAATGTT
AAGAGTTTGA CTGTAATTGC

          BamHI
          ~~~~~~
   281  GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT
GGCACTGGTA GTGATACGAA
          CTACCTAGGA GGCGCTGCTG TCGTCGTTTC TTTATTAGTG GCAGAAGGCA
CCGTGACCAT CACTATGCTT

   351  TCTACAACTC GATACTAACT ACACCCTCAC GCCTTTCGAC ACCCTACCAC
AATGCAACGG TTGTGAAGTA
          AGATGTTGAG CTATGATTGA TGTGGGAGTG CGGAAAGCTG TGGGATGGTG
TTACGTTGCC AACACTTCAT

   421  CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC
GCTTGTCAAA CAGCAGGTTA
          GTGCCACCTA TAATATAACC TACCCAGAGG CAGGTCCTGG TTCAGCTCAG
CGAACAGTTT GTCGTCCAAT

   491  GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC
TCCCTGGCGG CACTCACTGC
          CGGTCATAGG CCTGATGCGC GACTGGCACT GGCCGGTGMG GGAGCCGCGG
AGGGACCGCC GTGAGTGACG
```

$Fig. 43B$

561 CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG
AACCGCGCAG CGGCAATCAG
         GCGGGTCGAC AGACGCTGTA TGCTGTTGTA GGCGGACATG TGGAAGCCGC
TTGGCGCGTC GCCGTTAGTC


                                                       XhoI
                                                       ~~~~~~~
631 GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC
GACGCAGTAT TTCCGGGTCA
         CGGAAGCGCA GCATGTACTT GCTACGGAAG GTTCGGAGCT CGGGTCTATG
CTGCGTCATA AAGGCCCAGT


NcoI

~~~~~~
     701 CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC
GCCCATGGCG GTGTAGAGTA
         GAGTACGGTT GCTGCCGTAG GGTTTGGACG GGGGCCACCT CGTCCCCATG
CGGGTACCGC CACATCTCAT

     771 CTGGAGCGTT GATCCTTACA GCGCCCAGAA CACATTTGTC TGCACTGGGG
ATGAAGTGCA GTGCTGTGAG
         GACCTCGCAA CTAGGAATGT CGCGGGTCTT GTGTAAACAG ACGTGACCCC
TACTTCACGT CACGACACTC


SphI

~~~~~~
     841 GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT
GACGAGCGGC GCATGCACCT
         CGGGTCCCGC CTGTCCCACA CTTATTACGC GTGTGCTGAA TAAAACCCTA
CTGCTCGCCG CGTACGTGGA


                   NotI
                   ~~~~~~~~~
     911 GGCCGGTCGC GGCCGCGGAA ACCACTGAAG GATGAGCTGT AAAGAAGCAG
ATCGTTCAAA CATTTGGCAA
         CCGGCCAGCG CCGGCGCCTT TGGTGACTTC CTACTCGACA TTTCTTCGTC
TAGCAAGTTT GTAAACCGTT


     981 TAAAGTTTCT TAAGATTGAA TCCTGTTGCC GGTCTTGCGA TGATTATCAT
ATAATTTCTG TTGAATTACG
         ATTTCAAAGA ATTCTAACTT AGGACAACGG CCAGAACGCT ACTAATAGTA
TATTAAAGAC AACTTAATGC


     1051 TTAAGCATGT AATAATTAAC ATGTAATGCA TGACGTTATT TATGAGATGG
GTTTTTATGA TTAGAGTCCC
          AATTCGTACA TTATTAATTG TACATTACGT ACTGCAATAA ATACTCTACC
CAAAAATACT AATCTCAGGG


Fig. 43 C

BssHII

~~~~~~

~~~~~

1121   GCAATTATAC ATTTAATACG CGATAGAAAA CAAAATATAG CGCGCAAACT
AGGATAAATT ATCGCGCGCG
       CGTTAATATG TAAATTATGC GCTATCTTTT GTTTTATATC GCGCGTTTGA
TCCTATTTAA TAGCGCGCGC

                                    XbaI

                                 ~~~~~~

                 ClaI HindIII

               ~~~~~~~~~~~~~~

1191   GTGTCATCTA TGTTACTAGA TCGATAAGCT TCTAGAGCGG CCGGTGGAGC
TCCAATTCGC CCTATAGTGA
       CACAGTAGAT ACAATGATCT AGCTATTCGA AGATCTCGCC GGCCACCTCG
AGGTTAAGCG GGATATCACT

            BssHII

            ~~~~~~

1261   GTCGTATTAC GCGCGCTCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG
AAAACCCTGG CGTTACCCAA
       CAGCATAATG CGCGCGAGTG ACCGGCAGCA AAATGTTGCA GCACTGACCC
TTTTGGGACC GCAATGGGTT

1331   CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA
AGAGGCCCGC ACCGATCGCC
       GAATTAGCGG AACGTCGTGT AGGGGGAAAG CGGTCGACCG CATTATCGCT
TCTCCGGGCG TGGCTAGCGG

1401   CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGGACGC GCCCTGTAGC
GGCGCATTAA GCGCGGCGGG
       GAAGGGTTGT CAACGCGTCG GACTTACCGC TTACCCTGCG CGGGACATCG
CCGCGTAATT CGCGCCGCCC

1471   TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC
CCGCTCCTTT CGCTTTCTTC
       ACACCACCAA TGCGCGTCGC ACTGGCGATG TGAACGGTCG CGGGATCGCG
GGCGAGGAAA GCGAAAGAAG

1541   CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCGTCAAG CTCTAAATCG
GGGGCTCCCT TTAGGGTTCC
       GGAAGGAAAG AGCGGTGCAA GCGGCCGAAA GGGGCAGTTC GAGATTTAGC
CCCCGAGGGA AATCCCAAGG

1611   GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAACTTGA TTAGGGTGAT
GGTTCACGTA GTGGGCCATC
       CTAAATCACG AAATGCCGTG GAGCTGGGGT TTTTTGAACT AATCCCACTA
CCAAGTGCAT CACCCGGTAG

1681   GCCCTGATAG ACGGTTTTTC GCCCTTTGAC GTTGGAGTCC ACGTTCTTTA
ATAGTGGACT CTTGTTCCAA
       CGGGACTATC TGCCAAAAAG CGGGAAACTG CAACCTCAGG TGCAAGAAAT
TATCACCTGA GAACAAGGTT

Fig. 43D

1751   ACTGGAACAA CACTCAACCC TATCTCGGTC TATTCTTTTG ATTTATAAGG
GATTTTGCCG ATTTCGGCCT
        TGACCTTGTT GTGAGTTGGG ATAGAGCCAG ATAAGAAAC TAAATATTCC
CTAAAACGGC TAAAGCCGGA

1821   ATTGGTTAAA AAATGAGCTG ATTTAACAAA AATTTAACGC GAATTTTAAC
AAAATATTAA CGCTTACAAT
        TAACCAATTT TTTACTCGAC TAAATTGTTT TTAAATTGCG CTTAAAATTG
TTTTATAATT GCGAATGTTA

1891   TTAGGTGGCA CTTTTCGGGG AAATGTGCGC GGAACCCCTA TTTGTTTATT
TTTCTAAATA CATTCAAATA
        AATCCACCGT GAAAAGCCCC TTTACACGCG CCTTGGGGAT AAACAAATAA
AAAGATTTAT GTAAGTTTAT

1961   TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA
AAAAGGAAGA GTATGAGTAT
        ACATAGGCGA GTACTCTGTT ATTGGGACTA TTTACGAAGT TATTATAACT
TTTTCCTTCT CATACTCATA

2031   TCAACATTTC CGTGTCGCCC TTATTCCCTT TTTTGCGGCA TTTTGCCTTC
CTGTTTTTGC TCACCCAGAA
        AGTTGTAAAG GCACAGCGGG AATAAGGGAA AAAACGCCGT AAAACGGAAG
GACAAAAACG AGTGGGTCTT

2101   ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG
TTACATCGAA CTGGATCTCA
        TGCGACCACT TTCATTTTCT ACGACTTCTA GTCAACCCAC GTGCTCACCC
AATGTAGCTT GACCTAGAGT

2171   ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG TTTTCCAATG
ATGAGCACTT TTAAAGTTCT
        TGTCGCCATT CTAGGAACTC TCAAAAGCGG GGCTTCTTGC AAAAGGTTAC
TACTCGTGAA AATTTCAAGA

2241   GCTATGTGGC GCGGTATTAT CCCGTATTGA CGCCGGGCAA GAGCAACTCG
GTCGCCGCAT ACACTATTCT
        CGATACACCG CGCCATAATA GGGCATAACT GCGGCCCGTT CTCGTTGAGC
CAGCGGCGTA TGTGATAAGA

2311   CAGAATGACT TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA
TGGCATGACA GTAAGAGAAT
        GTCTTACTGA ACCAACTCAT GAGTGGTCAG TGTCTTTTCG TAGAATGCCT
ACCGTACTGT CATTCTCTTA

2381   TATGCAGTGC TGCCATAACC ATGAGTGATA ACACTGCGGC CAACTTACTT
CTGACAACGA TCGGAGGACC
        ATACGTCACG ACGGTATTGG TACTCACTAT TGTGACGCCG GTTGAATGAA
GACTGTTGCT AGCCTCCTGG

2451   GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACTCGCC
TTGATCGTTG GGAACCGGAG

Fig 43 E

```
        CTTCCTCGAT TGGCGAAAAA ACGTGTTGTA CCCCCTAGTA CATTGAGCGG
AACTAGCAAC CCTTGGCCTC

   2521 CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCACGA TGCCTGTAGC
AATGGCAACA ACGTTGCGCA
        GACTTACTTC GGTATGGTTT GCTGCTCGCA CTGTGGTGCT ACGGACATCG
TTACCGTTGT TGCAACGCGT

   2591 AACTATTAAC TGGCGAACTA CTTACTCTAG CTTCCCGGCA ACAATTAATA
GACTGGATGG AGGCGGATAA
        TTGATAATTG ACCGCTTGAT GAATGAGATC GAAGGGCCGT TGTTAATTAT
CTGACCTACC TCCGCCTATT

   2661 AGTTGCAGGA CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG
CTGATAAATC TGGAGCCGGT
        TCAACGTCCT GGTGAAGACG CGAGCCGGGA AGGCCGACCG ACCAAATAAC
GACTATTTAG ACCTCGGCCA

   2731 GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC
CTCCCGTATC GTAGTTATCT
        CTCGCACCCA GAGCGCCATA GTAACGTCGT GACCCCGGTC TACCATTCGG
GAGGGCATAG CATCAATAGA

   2801 ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT
GAGATAGGTG CCTCACTGAT
        TGTGCTGCCC CTCAGTCCGT TGATACCTAC TTGCTTTATC TGTCTAGCGA
CTCTATCCAC GGAGTGACTA

   2871 TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG
ATTTAAAACT TCATTTTTAA
        ATTCGTAACC ATTGACAGTC TGGTTCAAAT GAGTATATAT GAAATCTAAC
TAAATTTTGA AGTAAAAATT

   2941 TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT
CCCTTAACGT GAGTTTTCGT
        AAATTTTCCT AGATCCACTT CTAGGAAAAA CTATTAGAGT ACTGGTTTTA
GGGAATTGCA CTCAAAAGCA

   3011 TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT
CCTTTTTTTC TGCGCGTAAT
        AGGTGACTCG CAGTCTGGGG CATCTTTTCT AGTTTCCTAG AAGAACTCTA
GGAAAAAAAG ACGCGCATTA

   3081 CTGCTGCTTG CAAACAAAAA AACCACCGCT ACCAGCGGTG GTTTGTTTGC
CGGATCAAGA GCTACCAACT
        GACGACGAAC GTTTGTTTTT TTGGTGGCGA TGGTCGCCAC CAAACAAACG
GCCTAGTTCT CGATGGTTGA

   3151 CTTTTTCCGA AGGTAACTGG CTTCAGCAGA GCGCAGATAC CAAATACTGT
CCTTCTAGTG TAGCCGTAGT
        GAAAAAGGCT TCCATTGACC GAAGTCGTCT CGCGTCTATG GTTTATGACA
GGAAGATCAC ATCGGCATCA
```

Fig. 43 F

```
3221  TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG
CTAATCCTGT TACCAGTGGC
         ATCCGGTGGT GAAGTTCTTG AGACATCGTG GCGGATGTAT GGAGCGAGAC
GATTAGGACA ATGGTCACCG

3291  TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT
AGTTACCGGA TAAGGCGCAG
         ACGACGGTCA CCGCTATTCA GCACAGAATG GCCCAACCTG AGTTCTGCTA
TCAATGGCCT ATTCCGCGTC

3361  CGGTCGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC
GACCTACACC GAACTGAGAT
         GCCAGCCCGA CTTGCCCCCC AAGCACGTGT GTCGGGTCGA ACCTCGCTTG
CTGGATGTGG CTTGACTCTA

3431  ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG
GCGGACAGGT ATCCGGTAAG
         TGGATGTCGC ACTCGATACT CTTTCGCGGT GCGAAGGGCT TCCCTCTTTC
CGCCTGTCCA TAGGCCATTC

3501  CGGCAGGGTC GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGGAAACG
CCTGGTATCT TTATAGTCCT
         GCCGTCCCAG CCTTGTCCTC TCGCGTGCTC CCTCGAAGGT CCCCCTTTGC
GGACCATAGA AATATCAGGA

3571  GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTTGT GATGCTCGTC
AGGGGGGCGG AGCCTATGGA
         CAGCCCAAAG CGGTGGAGAC TGAACTCGCA GCTAAAAACA CTACGAGCAG
TCCCCCCGCC TCGGATACCT

3641  AAAACGCCAG CAACGCGGCC TTTTTACGGT TCCTGGCCTT TTGCTGGCCT
TTTGCTCACA TGTTCTTTCC
         TTTTGCGGTC GTTGCGCCGG AAAAATGCCA AGGACCGGAA AACGACCGGA
AAACGAGTGT ACAAGAAAGG

3711  TGCGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC TTTGAGTGAG
CTGATACCGC TCGCCGCAGC
         ACGCAATAGG GGACTAAGAC ACCTATTGGC ATAATGGCGG AAACTCACTC
GACTATGGCG AGCGGCGTCG

3781  CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCC
AATACGCAAA CCGCCTCTCC
         GCTTGCTGGC TCGCGTCGCT CAGTCACTCG CTCCTTCGCC TTCTCGCGGG
TTATGCGTTT GGCGGAGAGG

3851  CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG GTTTCCCGAC
TGGAAAGCGG GCAGTGAGCG
         GGCGCGCAAC CGGCTAAGTA ATTACGTCGA CCGTGCTGTC CAAAGGGCTG
ACCTTTCGCC CGTCACTCGC

3921  CAACGCAATT AATGTGAGTT AGCTCACTCA TTAGGCACCC CAGGCTTTAC
ACTTTATGCT TCCGGCTCGT
         GTTGCGTTAA TTACACTCAA TCGAGTGAGT AATCCGTGGG GTCCGAAATG
TGAAATACGA AGGCCGAGCA
```

Fig 43G

BssHII

~~
    3991   ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG GAAACAGCTA
TGACCATGAT TACGCCAAGC
          TACAACACAC CTTAACACTC GCCTATTGTT AAAGTGTGTC CTTTGTCGAT
ACTGGTACTA ATGCGGTTCG

       BssHII   EcoRI
       ~~~~   ~~~~~~~
    4061   GCGCAATTAA CCCTCACTAA AGGGAACAAA AGCTGGAATT CCACAATGAA
CAATAATAAG ATTAAAATAG
          CGCGTTAATT GGGAGTGATT TCCCTTGTTT TCGACCTTAA GGTGTTACTT
GTTATTATTC TAATTTTATC

    4131   CTTGCCCCCG TTGCAGCGAT GGGTATTTTT CTAGTAAAA TAAAAGATAA
ACTTAGACTC AAAACATTTA
          GAACGGGGGC AACGTCGCTA CCCATAAAAA AGATCATTTT ATTTTCTATT
TGAATCTGAG TTTTGTAAAT

    4201   CAAAAACAAC CCCTAAAGTC CTAAAGCCCA AAGTGCTATG CACGATCCAT
AGCAAGCCCA GCCCAACCCA
          GTTTTTGTTG GGGATTTCAG GATTTCGGGT TTCACGATAC GTGCTAGGTA
TCGTTCGGGT CGGGTTGGGT

    4271   ACCCAACCCA ACCCACCCCA GTGCAGCCAA CTGGCAAATA GTCTCCACCC
CCGGCACTAT CACCGTGAGT
          TGGGTTGGGT TGGGTGGGGT CACGTCGGTT GACCGTTTAT CAGAGGTGGG
GGCCGTGATA GTGGCACTCA

    4341   TGTCCGCACC ACCGCACGTC TCGCAGCCAA AAAAAAAAA AGAAAGAAAA
AAAAGAAAAA GAAAAACAGC
          ACAGGCGTGG TGGCGTGCAG AGCGTCGGTT TTTTTTTTTT TCTTTCTTTT
TTTTCTTTTT CTTTTTGTCG

    4411   AGGTGGGTCC GGGTCGTGGG GGCCGGAAAA GCGAGGAGGA TCGCGAGCAG
CGACGAGGCC CGGCCCTCCC
          TCCACCCAGG CCCAGCACCC CCGGCCTTTT CGCTCCTCCT AGCGCTCGTC
GCTGCTCCGG GCCGGGAGGG

    4481   TCCGCTTCCA AGAAACGCC CCCCATCGCC ACTATATACA TACCCCCCCC
TCTCCTCCCA TCCCCCCAAC
          AGGCGAAGGT TTCTTTGCGG GGGGTAGCGG TGATATATGT ATGGGGGGGG
AGAGGAGGGT AGGGGGGTTG

    4551   CCTACCACCA CCACCACCAC CACCTCCTCC CCCCTCGCTG CCGGACGACG
AGCTCCTCCC CCCTCCCCCT
          GGATGGTGGT GGTGGTGGTG GTGGAGGAGG GGGGAGCGAC GGCCTGCTGC
TCGAGGAGGG GGGAGGGGGA

    4621   CCGCCGCCGC CGGTAACCAC CCCGCCCCTC TCCTCTTTCT TTCTCCGTTT
TTTTTTTCGT CTCGGTCTCG

Fig 43H

```
        GGCGGCGGCG GCCATTGGTG GGGCGGGGAG AGGAGAAAGA AAGAGGCAAA
AAAAAAAGCA GAGCCAGAGC

   4691  ATCTTTGGCC TTGGTAGTTT GGGTGGGCGA GAGCGGCTTC GTCGCCCAGA
TCGGTGCGCG GGAGGGGCGG
        TAGAAACCGG AACCATCAAA CCCACCCGCT CTCGCCGAAG CAGCGGGTCT
AGCCACGCGC CCTCCCCGCC

                                                                BamHI

BglII
                                                                ~~~~~~~
~
   4761  GATCTCGCGG CTGGCGTCTC CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG
GAATGGGGCT CTCGGATGTA
        CTAGAGCGCC GACCGCAGAG GCCCGCACTC AGCCGGGCCT AGGAGCGCCC
CTTACCCCGA GAGCCTACAT

        BglII
        ~~~~~
   4831  GATCTTCTTT CTTTCTTCTT TTTGTGGTAG AATTTGAATC CCTCAGCATT
GTTCATCGGT AGTTTTTCTT
        CTAGAAGAAA GAAAGAAGAA AAACACCATC TTAAACTTAG GGAGTCGTAA
CAAGTAGCCA TCAAAAAGAA

   4901  TTCATGATTT GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGGTAG
        AAGTACTAAA CACTGTTTAC GTCGGAGCAC GCCTCGAAAA AACATCCATC
```

Fig 43 I

Bss HII (4938)   *Eco* RI (1)

Actin promoter

Bam HI (703)

Bgl II (735)

Hin dIII (856)

ori   PPI (APO)

Nco I (938)

Not I (1007)

pJQ4
4974 bp

Sal I (1107)

Acc I (1108)

Bam HI (1163)

SIG-VAC-FAE-F/S

Xho I (1545)

Nco I (1632)

amp

Sph I (1784)

f1 ori

Not I (1799)

Bss HII (2039)

Bss HII (2063)

Cla I (2090)

Hin dIII (2095)

Xba I (2100)

Bss HII (2150)

Fig. 44 A

Sequence for pJQ4

```
        EcoRI
        ~~~~~
     1  AATTCCACAA TGAACAATAA TAAGATTAAA ATAGCTTGCC CCCGTTGCAG
CGATGGGTAT TTTTTCTAGT
        TTAAGGTGTT ACTTGTTATT ATTCTAATTT TATCGAACGG GGGCAACGTC
GCTACCCATA AAAAAGATCA
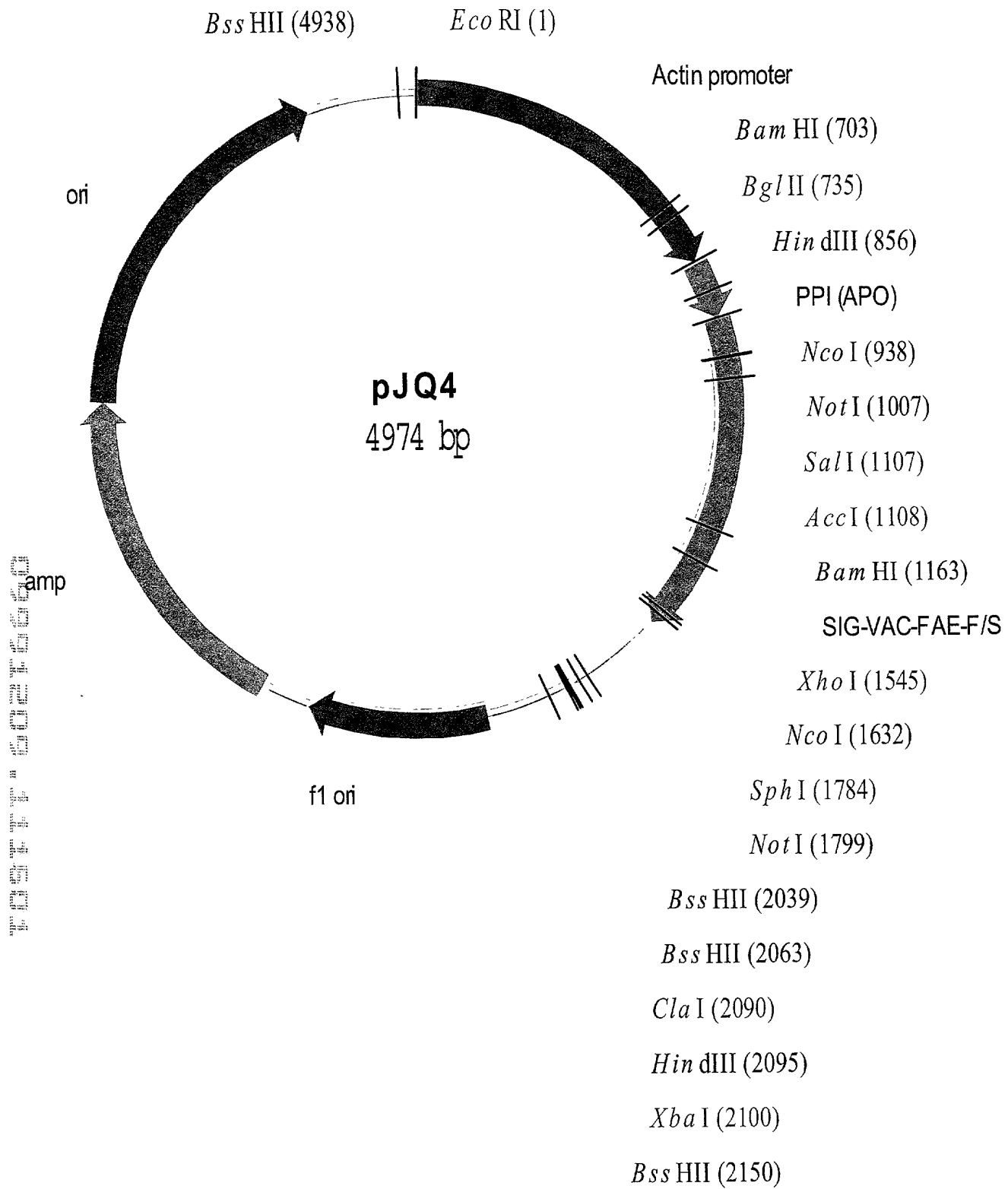
    71  AAAATAAAAG ATAAACTTAG ACTCAAAACA TTTACAAAAA CAACCCCTAA
AGTCCTAAAG CCCAAAGTGC
        TTTTATTTTC TATTTGAATC TGAGTTTTGT AAATGTTTTT GTTGGGGATT
TCAGGATTTC GGGTTTCACG

   141  TATGCACGAT CCATAGCAAG CCCAGCCCAA CCCAACCCAA CCCAACCCAC
CCCAGTGCAG CCAACTGGCA
        ATACGTGCTA GGTATCGTTC GGGTCGGGTT GGGTTGGGTT GGGTTGGGTG
GGGTCACGTC GGTTGACCGT

   211  AATAGTCTCC ACCCCCGGCA CTATCACCGT GAGTTGTCCG CACCACCGCA
CGTCTCGCAG CCAAAAAAAA
        TTATCAGAGG TGGGGGCCGT GATAGTGGCA CTCAACAGGC GTGGTGGCGT
GCAGAGCGTC GGTTTTTTTT

   281  AAAAGAAAG AAAAAAAAGA AAAAGAAAAA CAGCAGGTGG GTCCGGGTCG
TGGGGGCCGG AAAAGCGAGG
        TTTTTCTTTC TTTTTTTTCT TTTTCTTTTT GTCGTCCACC CAGGCCCAGC
ACCCCCGGCC TTTTCGCTCC

   351  AGGATCGCGA GCAGCGACGA GGCCCGGCCC TCCCTCCGCT TCCAAAGAAA
CGCCCCCCAT CGCCACTATA
        TCCTAGCGCT CGTCGCTGCT CCGGGCCGGG AGGGAGGCGA AGGTTTCTTT
GCGGGGGGTA GCGGTGATAT

   421  TACATACCCC CCCCTCTCCT CCCATCCCCC CAACCCTACC ACCACCACCA
CCACCACCTC CTCCCCCCTC
        ATGTATGGGG GGGGAGAGGA GGGTAGGGGG GTTGGGATGG TGGTGGTGGT
GGTGGTGGAG GAGGGGGGAG

   491  GCTGCCGGAC GACGAGCTCC TCCCCCCTCC CCCTCCGCCG CCGCCGGTAA
CCACCCCGCC CCTCTCCTCT
        CGACGGCCTG CTGCTCGAGG AGGGGGGAGG GGGAGGCGGC GGCGGCCATT
GGTGGGGCGG GGAGAGGAGA

   561  TTCTTTCTCC GTTTTTTTTT TCGTCTCGGT CTCGATCTTT GGCCTTGGTA
GTTTGGGTGG GCGAGAGCGG
        AAGAAAGAGG CAAAAAAAAA AGCAGAGCCA GAGCTAGAAA CCGGAACCAT
CAAACCCACC CGCTCTCGCC

   631  CTTCGTCGCC CAGATCGGTG CGCGGGAGGG GCGGGATCTC GCGGCTGGCG
TCTCCGGGCG TGAGTCGGCC
```

Fig. 44B

```
          GAAGCAGCGG GTCTAGCCAC GCGCCCTCCC CGCCCTAGAG CGCCGACCGC
       AGAGGCCCGC ACTCAGCCGG
```

```
                     BamHI                              BglII
                     ~~~~~~                              ~~~~~~
           701   CGGATCCTCG CGGGGAATGG GGCTCTCGGA TGTAGATCTT CTTTCTTTCT
       TCTTTTTGTG GTAGAATTTG
                 GCCTAGGAGC GCCCCTTACC CCGAGAGCCT ACATCTAGAA GAAAGAAAGA
       AGAAAAACAC CATCTTAAAC
```

```
           771   AATCCCTCAG CATTGTTCAT CGGTAGTTTT TCTTTTCATG ATTTGTGACA
       AATGCAGCCT CGTGCGGAGC
                 TTAGGGAGTC GTAACAAGTA GCCATCAAAA AGAAAAGTAC TAAACACTGT
       TTACGTCGGA GCACGCCTCG
```

```
                           HindIII
                           ~~~~~~
           841   TTTTTTGTAG GTAGAAGCTT ACMATGGMCG TGCACAAGGA GGTSAACTTC
       GTSGCCTACC TCCTGATCGT
                 AAAAAACATC CATCTTCGAA TGKTACCKGC ACGTGTTCCT CCASTTGAAG
       CASCGGATGG AGGACTAGCA
```

```
                                      NcoI
                                      ~~~~~~~
           911   SCTCGGCCTC CTCTTGCTCG TSTCCGCCAT GGAGCACGTG GACGCCAAGG
       CCTGCACCCK CGAGTGCGGC
                 SGAGCCGGAG GAGAACGAGC ASAGGCGGTA CCTCGTGCAC CTGCGGTTCC
       GGACGTGGGM GCTCACGCCG
```

```
                                      NotI
                                      ~~~~~~~~~
           981   AACCTCGGCT TCGGCATCTG CCCGGCGGCC GCCTCCACGC AGGGCATCTC
       CGAAGACCTC TACAGCCGTT
                 TTGGAGCCGA AGCCGTAGAC GGGCCGCCGG CGGAGGTGCG TCCCGTAGAG
       GCTTCTGGAG ATGTCGGCAA
```

```
       SalI

       ~~~~~~~

       AccI

       ~~~~~~~
          1051   TAGTCGAAAT GGCCACTATC TCCCAAGCTG CCTACGCCGA CCTGTGCAAC
       ATTCCGTCGA CTATTATCAA
                 ATCAGCTTTA CCGGTGATAG AGGGTTCGAC GGATGCGGCT GGACACGTTG
       TAAGGCAGCT GATAATAGTT
```

```
                                                          BamHI
                                                          ~~~~~~
          1121   GGGAGAGAAA ATTTACAATT CTCAAACTGA CATTAACGGA TGGATCCTCC
       GCGACGACAG CAGCAAAGAA
```

Fig. 44 G

CCCTCTCTTT TAAATGTTAA GAGTTTGACT GTAATTGCCT ACCTAGGAGG
CGCTGCTGTC GTCGTTTCTT

    1191 ATAATCACCG TCTTCCGTGG CACTGGTAGT GATACGAATC TACAACTCGA
TACTAACTAC ACCCTCACGC
          TATTAGTGGC AGAAGGCACC GTGACCATCA CTATGCTTAG ATGTTGAGCT
ATGATTGATG TGGGAGTGCG

    1261 CTTTCGACAC CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT
TATATTGGAT GGGTCTCCGT
          GAAAGCTGTG GGATGGTGTT ACGTTGCCAA CACTTCATGT GCCACCTATA
ATATAACCTA CCCAGAGGCA

    1331 CCAGGACCAA GTCGAGTCGC TTGTCAAACA GCAGGTTAGC CAGTATCCGG
ACTACGCGCT GACCGTGACC
          GGTCCTGGTT CAGCTCAGCG AACAGTTTGT CGTCCAATCG GTCATAGGCC
TGATGCGCGA CTGGCACTGG

    1401 GGCCACKCCC TCGGCGCCTC CCTGGCGGCA CTCACTGCCG CCCAGCTGTC
TGCGACATAC GACAACATCC
          CCGGTGMGGG AGCCGCGGAG GGACCGCCGT GAGTGACGGC GGGTCGACAG
ACGCTGTATG CTGTTGTAGG

    1471 GCCTGTACAC CTTCGGCGAA CCGCGCAGCG GCAATCAGGC CTTCGCGTCG
TACATGAACG ATGCCTTCCA
          CGGACATGTG GAAGCCGCTT GGCGCGTCGC CGTTAGTCCG GAAGCGCAGC
ATGTACTTGC TACGGAAGGT

                         XhoI
                        ~~~~~~
    1541 AGCCTCGAGC CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAACG
ACGGCATCCC AAACCTGCCC
          TCGGAGCTCG GGTCTATGCT GCGTCATAAA GGCCCAGTGA GTACGGTTGC
TGCCGTAGGG TTTGGACGGG

                              NcoI
                             ~~~~~~
    1611 CCGGTGGAGC AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA
TCCTTACAGC GCCCAGAACA
          GGCCACCTCG TCCCCATGCG GGTACCGCCA CATCTCATGA CCTCGCAACT
AGGAATGTCG CGGGTCTTGT

    1681 CATTTGTCTG CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA
CAGGGTGTGA ATAATGCGCA
          GTAAACAGAC GTGACCCCTA CTTCACGTCA CGACACTCCG GGTCCCGCCT
GTCCCACACT TATTACGCGT

                                      SphI   NotI
                                     ~~~~~~~   ~~~~~~~~~
    1751 CACGACTTAT TTTGGGATGA CGAGCGGCGC ATGCACCTGG CCGGTCGCGG
CCGCGGAAAC CACTGAAGGA
          GTGCTGAATA AAACCCTACT GCTCGCCGCG TACGTGGACC GGCCAGCGCC
GGCGCCTTTG GTGACTTCCT

Fig. 44 D

1821 TGAGCTGTAA AGAAGCAGAT CGTTCAAACA TTTGGCAATA AAGTTTCTTA
AGATTGAATC CTGTTGCCGG
        ACTCGACATT TCTTCGTCTA GCAAGTTTGT AAACCGTTAT TTCAAAGAAT
TCTAACTTAG GACAACGGCC

1891 TCTTGCGATG ATTATCATAT AATTTCTGTT GAATTACGTT AAGCATGTAA
TAATTAACAT GTAATGCATG
        AGAACGCTAC TAATAGTATA TTAAAGACAA CTTAATGCAA TTCGTACATT
ATTAATTGTA CATTACGTAC

1961 ACGTTATTTA TGAGATGGGT TTTTATGATT AGAGTCCCGC AATTATACAT
TTAATACGCG ATAGAAAACA
        TGCAATAAAT ACTCTACCCA AAAATACTAA TCTCAGGGCG TTAATATGTA
AATTATGCGC TATCTTTTGT

XbaI

~~
                    BssHII   BssHII
ClaI HindIII
                    ~~~~~~~   ~~~~~~

~~~~~~~~~~~~~~~
    2031 AAATATAGCG CGCAAACTAG GATAAATTAT CGCGCGCGGT GTCATCTATG
TTACTAGATC GATAAGCTTC
        TTTATATCGC GCGTTTGATC CTATTTAATA GCGCGCGCCA CAGTAGATAC
AATGATCTAG CTATTCGAAG

        XbaI   BssHII
        ~~~~   ~~~~~~~
    2101 TAGAGCGGCC GGTGGAGCTC CAATTCGCCC TATAGTGAGT CGTATTACGC
GCGCTCACTG GCCGTCGTTT
        ATCTCGCCGG CCACCTCGAG GTTAAGCGGG ATATCACTCA GCATAATGCG
CGCGAGTGAC CGGCAGCAAA

    2171 TACAACGTCG TGACTGGGAA AACCCTGGCG TTACCCAACT TAATCGCCTT
GCAGCACATC CCCCTTTCGC
        ATGTTGCAGC ACTGACCCTT TTGGGACCGC AATGGGTTGA ATTAGCGGAA
CGTCGTGTAG GGGGAAAGCG

    2241 CAGCTGGCGT AATAGCGAAG AGGCCCGCAC CGATCGCCCT TCCCAACAGT
TGCGCAGCCT GAATGGCGAA
        GTCGACCGCA TTATCGCTTC TCCGGGCGTG GCTAGCGGGA AGGGTTGTCA
ACGCGTCGGA CTTACCGCTT

    2311 TGGGACGCGC CCTGTAGCGG CGCATTAAGC GCGGCGGGTG TGGTGGTTAC
GCGCAGCGTG ACCGCTACAC
        ACCCTGCGCG GGACATCGCC GCGTAATTCG CGCCGCCCAC ACCACCAATG
CGCGTCGCAC TGGCGATGTG

    2381 TTGCCAGCGC CCTAGCGCCC GCTCCTTTCG CTTTCTTCCC TTCCTTTCTC
GCCACGTTCG CCGGCTTTCC
        AACGGTCGCG GGATCGCGGG CGAGGAAAGC GAAAGAAGGG AAGGAAAGAG
CGGTGCAAGC GGCCGAAAGG

Fig. 44 E

2451   CCGTCAAGCT CTAAATCGGG GGCTCCCTTT AGGGTTCCGA TTTAGTGCTT
TACGGCACCT CGACCCCAAA
        GGCAGTTCGA GATTTAGCCC CCGAGGGAAA TCCCAAGGCT AAATCACGAA
ATGCCGTGGA GCTGGGGTTT

2521   AAACTTGATT AGGGTGATGG TTCACGTAGT GGGCCATCGC CCTGATAGAC
GGTTTTTCGC CCTTTGACGT
        TTTGAACTAA TCCCACTACC AAGTGCATCA CCCGGTAGCG GGACTATCTG
CCAAAAAGCG GGAAACTGCA

2591   TGGAGTCCAC GTTCTTTAAT AGTGGACTCT TGTTCCAAAC TGGAACAACA
CTCAACCCTA TCTCGGTCTA
        ACCTCAGGTG CAAGAAATTA TCACCTGAGA ACAAGGTTTG ACCTTGTTGT
GAGTTGGGAT AGAGCCAGAT

2661   TTCTTTTGAT TTATAAGGGA TTTTGCCGAT TTCGGCCTAT TGGTTAAAAA
ATGAGCTGAT TTAACAAAAA
        AAGAAAACTA AATATTCCCT AAAACGGCTA AAGCCGGATA ACCAATTTTT
TACTCGACTA AATTGTTTTT

2731   TTTAACGCGA ATTTTAACAA AATATTAACG CTTACAATTT AGGTGGCACT
TTTCGGGGAA ATGTGCGCGG
        AAATTGCGCT TAAAATTGTT TTATAATTGC GAATGTTAAA TCCACCGTGA
AAAGCCCCTT TACACGCGCC

2801   AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA
TGAGACAATA ACCCTGATAA
        TTGGGGATAA ACAAATAAAA AGATTTATGT AAGTTTATAC ATAGGCGAGT
ACTCTGTTAT TGGGACTATT

2871   ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG
TGTCGCCCTT ATTCCCTTTT
        TACGAAGTTA TTATAACTTT TTCCTTCTCA TACTCATAAG TTGTAAAGGC
ACAGCGGGAA TAAGGGAAAA

2941   TTGCGGCATT TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA
GTAAAAGATG CTGAAGATCA
        AACGCCGTAA AACGGAAGGA CAAAAACGAG TGGGTCTTTG CGACCACTTT
CATTTTCTAC GACTTCTAGT

3011   GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA
TCCTTGAGAG TTTTCGCCCC
        CAACCCACGT GCTCACCCAA TGTAGCTTGA CCTAGAGTTG TCGCCATTCT
AGGAACTCTC AAAAGCGGGG

3081   GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC
GGTATTATCC CGTATTGACG
        CTTCTTGCAA AAGGTTACTA CTCGTGAAAA TTTCAAGACG ATACACCGCG
CCATAATAGG GCATAACTGC

3151   CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
GTTGAGTACT CACCAGTCAC

Fig. 44F

```
        GGCCCGTTCT CGTTGAGCCA GCGGCGTATG TGATAAGAGT CTTACTGAAC
CAACTCATGA GTGGTCAGTG

   3221  AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG
CCATAACCAT GAGTGATAAC
        TCTTTTCGTA GAATGCCTAC CGTACTGTCA TTCTCTTAAT ACGTCACGAC
GGTATTGGTA CTCACTATTG

   3291  ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC
CGCTTTTTTG CACAACATGG
        TGACGCCGGT TGAATGAAGA CTGTTGCTAG CCTCCTGGCT TCCTCGATTG
GCGAAAAAAC GTGTTGTACC

   3361  GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC
ATACCAAACG ACGAGCGTGA
        CCCTAGTACA TTGAGCGGAA CTAGCAACCC TTGGCCTCGA CTTACTTCGG
TATGGTTTGC TGCTCGCACT

   3431  CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAAA CTATTAACTG
GCGAACTACT TACTCTAGCT
        GTGGTGCTAC GGACATCGTT ACCGTTGTTG CAACGCGTTT GATAATTGAC
CGCTTGATGA ATGAGATCGA

   3501  TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC
ACTTCTGCGC TCGGCCCTTC
        AGGGCCGTTG TTAATTATCT GACCTACCTC CGCCTATTTC AACGTCCTGG
TGAAGACGCG AGCCGGGAAG

   3571  CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT
CGCGGTATCA TTGCAGCACT
        GCCGACCGAC CAAATAACGA CTATTTAGAC CTCGGCCACT CGCACCCAGA
GCGCCATAGT AACGTCGTGA

   3641  GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA
GTCAGGCAAC TATGGATGAA
        CCCCGGTCTA CCATTCGGGA GGGCATAGCA TCAATAGATG TGCTGCCCCT
CAGTCCGTTG ATACCTACTT

   3711  CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTGGTA
ACTGTCAGAC CAAGTTTACT
        GCTTTATCTG TCTAGCGACT CTATCCACGG AGTGACTAAT TCGTAACCAT
TGACAGTCTG GTTCAAATGA

   3781  CATATATACT TTAGATTGAT TTAAAACTTC ATTTTTAATT TAAAAGGATC
TAGGTGAAGA TCCTTTTTGA
        GTATATATGA AATCTAACTA AATTTTGAAG TAAAAATTAA ATTTTCCTAG
ATCCACTTCT AGGAAAAACT

   3851  TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT
CAGACCCCGT AGAAAAGATC
        ATTAGAGTAC TGGTTTTAGG GAATTGCACT CAAAAGCAAG GTGACTCGCA
GTCTGGGGCA TCTTTTCTAG
```

Fig. 44 G

3921 AAAGGATCTT CTTGAGATCC TTTTTTTCTG CGCGTAATCT GCTGCTTGCA
AACAAAAAAA CCACCGCTAC
        TTTCCTAGAA GAACTCTAGG AAAAAAAGAC GCGCATTAGA CGACGAACGT
TTGTTTTTTT GGTGGCGATG

3991 CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG
GTAACTGGCT TCAGCAGAGC
        GTCGCCACCA AACAAACGGC CTAGTTCTCG ATGGTTGAGA AAAAGGCTTC
CATTGACCGA AGTCGTCTCG

4061 GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA GGCCACCACT
TCAAGAACTC TGTAGCACCG
        CGTCTATGGT TTATGACAGG AAGATCACAT CGGCATCAAT CCGGTGGTGA
AGTTCTTGAG ACATCGTGGC

4131 CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG
CGATAAGTCG TGTCTTACCG
        GGATGTATGG AGCGAGACGA TTAGGACAAT GGTCACCGAC GACGGTCACC
GCTATTCAGC ACAGAATGGC

4201 GGTTGGACTC AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA
ACGGGGGGTT CGTGCACACA
        CCAACCTGAG TTCTGCTATC AATGGCCTAT TCCGCGTCGC CAGCCCGACT
TGCCCCCCAA GCACGTGTGT

4271 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG
AGCTATGAGA AAGCGCCACG
        CGGGTCGAAC CTCGCTTGCT GGATGTGGCT TGACTCTATG GATGTCGCAC
TCGATACTCT TTCGCGGTGC

4341 CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG
AACAGGAGAG CGCACGAGGG
        GAAGGGCTTC CCTCTTTCCG CCTGTCCATA GGCCATTCGC CGTCCCAGCC
TTGTCCTCTC GCGTGCTCCC

4411 AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTCGC
CACCTCTGAC TTGAGCGTCG
        TCGAAGGTCC CCCTTTGCGG ACCATAGAAA TATCAGGACA GCCCAAAGCG
GTGGAGACTG AACTCGCAGC

4481 ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA AACGCCAGCA
ACGCGGCCTT TTTACGGTTC
        TAAAAACACT ACGAGCAGTC CCCCCGCCTC GGATACCTTT TTGCGGTCGT
TGCGCCGGAA AAATGCCAAG

4551 CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTTCCTG CGTTATCCCC
TGATTCTGTG GATAACCGTA
        GACCGGAAAA CGACCGGAAA ACGAGTGTAC AAGAAAGGAC GCAATAGGGG
ACTAAGACAC CTATTGGCAT

4621 TTACCGCCTT TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG
CGCAGCGAGT CAGTGAGCGA
        AATGGCGGAA ACTCACTCGA CTATGGCGAG CGGCGTCGGC TTGCTGGCTC
GCGTCGCTCA GTCACTCGCT

Fig. 44 H

```
4691   GGAAGCGGAA GAGCGCCCAA TACGCAAACC GCCTCTCCCC GCGCGTTGGC
CGATTCATTA ATGCAGCTGG
       CCTTCGCCTT CTCGCGGGTT ATGCGTTTGG CGGAGAGGGG CGCGCAACCG
GCTAAGTAAT TACGTCGACC

4761   CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA
TGTGAGTTAG CTCACTCATT
       GTGCTGTCCA AAGGGCTGAC CTTTCGCCCG TCACTCGCGT TGCGTTAATT
ACACTCAATC GAGTGAGTAA

4831   AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA
ATTGTGAGCG GATAACAATT
       TCCGTGGGGT CCGAAATGTG AAATACGAAG GCCGAGCATA CAACACACCT
TAACACTCGC CTATTGTTAA

                                              BssHII
                                              ~~~~~~~
4901   TCACACAGGA AACAGCTATG ACCATGATTA CGCCAAGCGC GCAATTAACC
CTCACTAAAG GGAACAAAAG
       AGTGTGTCCT TTGTCGATAC TGGTACTAAT GCGGTTCGCG CGTTAATTGG
GAGTGATTTC CCTTGTTTTC

       EcoR

4971   CTGG
       GACC
```

Fig. 44 I

Hin dIII (2)

Bgl II (5045)

Nco I (9)

Bam HI (5013)

sGFP

Rice actin

Not I (733)

Sma I (741)

Eco RI (4311)

Pst I (749)

Nos?

Xho I (3935)

pPQ10.1
5164 bp

Bss HII (984)

Bss HII (1008)

Kpn I (3924)

Cla I (1035)

Bss HII (3883)

Hin dIII (1040)

Xba I (1045)

Bss HII (1095)

f1 ori

ori

amp

Fig. 45 A

Sequence for pPQ10.1

```
            HindIII NcoI
            ~~~~~~ ~~~~~~~
      1    AAGCTTACCA TGGTGAGCAA GGGCGAGGAG CTGTTCACCG GGGTGGTGCC
CATCCTGGTC GAGCTGGACG
           TTCGAATGGT ACCACTCGTT CCCGCTCCTC GACAAGTGGC CCCACCACGG
GTAGGACCAG CTCGACCTGC

     71    GCGACGTGAA CGGCCACAAG TTCAGCGTGT CCGGCGAGGG CGAGGGCGAT
GCCACCTACG GCAAGCTGAC
           CGCTGCACTT GCCGGTGTTC AAGTCGCACA GGCCGCTCCC GCTCCCGCTA
CGGTGGATGC CGTTCGACTG

    141    CCTGAAGTTC ATCTGCACCA CCGGCAAGCT GCCCGTGCCC TGGCCCACCC
TCGTGACCAC CTTCACCTAC
           GGACTTCAAG TAGACGTGGT GGCCGTTCGA CGGGCACGGG ACCGGGTGGG
AGCACTGGTG GAAGTGGATG

    211    GGCGTGCAGT GCTTCAGCCG CTACCCCGAC CACATGAAGC AGCACGACTT
CTTCAAGTCC GCCATGCCCG
           CCGCACGTCA CGAAGTCGGC GATGGGGCTG GTGTACTTCG TCGTGCTGAA
GAAGTTCAGG CGGTACGGGC

    281    AAGGCTACGT CCAGGAGCGC ACCATCTTCT TCAAGGACGA CGGCAACTAC
AAGACCCGCG CCGAGGTGAA
           TTCCGATGCA GGTCCTCGCG TGGTAGAAGA AGTTCCTGCT GCCGTTGATG
TTCTGGGCGC GGCTCCACTT

    351    GTTCGAGGGC GACACCCTGG TGAACCGCAT CGAGCTGAAG GGCATCGACT
TCAAGGAGGA CGGCAACATC
           CAAGCTCCCG CTGTGGGACC ACTTGGCGTA GCTCGACTTC CCGTAGCTGA
AGTTCCTCCT GCCGTTGTAG

    421    CTGGGGCACA AGCTGGAGTA CAACTACAAC AGCCACAACG TCTATATCAT
GGCCGACAAG CAGAAGAACG
           GACCCCGTGT TCGACCTCAT GTTGATGTTG TCGGTGTTGC AGATATAGTA
CCGGCTGTTC GTCTTCTTGC

    491    GCATCAAGGT GAACTTCAAG ATCCGCCACA ACATCGAGGA CGGCAGCGTG
CAGCTCGCCG ACCACTACCA
           CGTAGTTCCA CTTGAAGTTC TAGGCGGTGT TGTAGCTCCT GCCGTCGCAC
GTCGAGCGGC TGGTGATGGT

    561    GCAGAACACC CCCATCGGCG ACGGCCCCGT GCTGCTGCCC GACAACCACT
ACCTGAGCAC CCAGTCCGCC
           CGTCTTGTGG GGGTAGCCGC TGCCGGGGCA CGACGACGGG CTGTTGGTGA
TGGACTCGTG GGTCAGGCGG

    631    CTGAGCAAAG ACCCCAACGA GAAGCGCGAT CACATGGTCC TGCTGGAGTT
CGTGACCGCC GCCGGGATCA
```

Fig. 45B

```
          CCGCATTATC GCTTCTCCGG GCGTGGCTAG CGGGAAGGGT TGTCAACGCG
TCGGACTTAC CGCTTACCCT

   1261   CGCGCCCTGT AGCGGCGCAT TAAGCGCGGC GGGTGTGGTG GTTACGCGCA
GCGTGACCGC TACACTTGCC
          GCGCGGGACA TCGCCGCGTA ATTCGCGCCG CCCACACCAC CAATGCGCGT
CGCACTGGCG ATGTGAACGG

   1331   AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC TTCCCTTCCT TTCTCGCCAC
GTTCGCCGGC TTTCCCCGTC
          TCGCGGGATC GCGGGCGAGG AAAGCGAAAG AAGGGAAGGA AAGAGCGGTG
CAAGCGGCCG AAAGGGGCAG

   1401   AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT TCCGATTTAG TGCTTTACGG
CACCTCGACC CCAAAAAACT
          TTCGAGATTT AGCCCCCGAG GGAAATCCCA AGGCTAAATC ACGAAATGCC
GTGGAGCTGG GGTTTTTTGA

   1471   TGATTAGGGT GATGGTTCAC GTAGTGGGCC ATCGCCCTGA TAGACGGTTT
TTCGCCCTTT GACGTTGGAG
          ACTAATCCCA CTACCAAGTG CATCACCCGG TAGCGGGACT ATCTGCCAAA
AAGCGGGAAA CTGCAACCTC

   1541   TCCACGTTCT TTAATAGTGG ACTCTTGTTC CAAACTGGAA CAACACTCAA
CCCTATCTCG GTCTATTCTT
          AGGTGCAAGA AATTATCACC TGAGAACAAG GTTTGACCTT GTTGTGAGTT
GGGATAGAGC CAGATAAGAA

   1611   TTGATTTATA AGGGATTTTG CCGATTTCGG CCTATTGGTT AAAAAATGAG
CTGATTTAAC AAAAATTTAA
          AACTAAATAT TCCCTAAAAC GGCTAAAGCC GGATAACCAA TTTTTTACTC
GACTAAATTG TTTTTAAATT

   1681   CGCGAATTTT AACAAAATAT TAACGCTTAC AATTTAGGTG GCACTTTTCG
GGGAAATGTG CGCGGAACCC
          GCGCTTAAAA TTGTTTTATA ATTGCGAATG TTAAATCCAC CGTGAAAAGC
CCCTTTACAC GCGCCTTGGG

   1751   CTATTTGTTT ATTTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA
CAATAACCCT GATAAATGCT
          GATAAACAAA TAAAAAGATT TATGTAAGTT TATACATAGG CGAGTACTCT
GTTATTGGGA CTATTTACGA

   1821   TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG
CCCTTATTCC CTTTTTTGCG
          AGTTATTATA ACTTTTTCCT TCTCATACTC ATAAGTTGTA AAGGCACAGC
GGGAATAAGG GAAAAAACGC

   1891   GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA
AGATGCTGAA GATCAGTTGG
          CGTAAAACGG AAGGACAAAA ACGAGTGGGT CTTTGCGACC ACTTTCATTT
TCTACGACTT CTAGTCAACC
```

Fig. 45 C

```
1961  GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT
      GAGAGTTTTC GCCCCGAAGA
          CACGTGCTCA CCCAATGTAG CTTGACCTAG AGTTGTCGCC ATTCTAGGAA
      CTCTCAAAAG CGGGGCTTCT

2031  ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGCGCGGTAT
      TATCCCGTAT TGACGCCGGG
          TGCAAAAGGT TACTACTCGT GAAAATTTCA AGACGATACA CCGCGCCATA
      ATAGGGCATA ACTGCGGCCC

2101  CAAGAGCAAC TCGGTCGCCG CATACACTAT TCTCAGAATG ACTTGGTTGA
      GTACTCACCA GTCACAGAAA
          GTTCTCGTTG AGCCAGCGGC GTATGTGATA AGAGTCTTAC TGAACCAACT
      CATGAGTGGT CAGTGTCTTT

2171  AGCATCTTAC GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA
      ACCATGAGTG ATAACACTGC
          TCGTAGAATG CCTACCGTAC TGTCATTCTC TTAATACGTC ACGACGGTAT
      TGGTACTCAC TATTGTGACG

2241  GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT
      TTTTGCACAA CATGGGGGAT
          CCGGTTGAAT GAAGACTGTT GCTAGCCTCC TGGCTTCCTC GATTGGCGAA
      AAAACGTGTT GTACCCCCTA

2311  CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC
      AAACGACGAG CGTGACACCA
          GTACATTGAG CGGAACTAGC AACCCTTGGC CTCGACTTAC TTCGGTATGG
      TTTGCTGCTC GCACTGTGGT

2381  CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA
      CTACTTACTC TAGCTTCCCG
          GCTACGGACA TCGTTACCGT TGTTGCAACG CGTTTGATAA TTGACCGCTT
      GATGAATGAG ATCGAAGGGC

2451  GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA GGACCACTTC
      TGCGCTCGGC CCTTCCGGCT
          CGTTGTTAAT TATCTGACCT ACCTCCGCCT ATTTCAACGT CCTGGTGAAG
      ACGCGAGCCG GGAAGGCCGA

2521  GGCTGGTTTA TTGCTGATAA ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG
      TATCATTGCA GCACTGGGGC
          CCGACCAAAT AACGACTATT TAGACCTCGG CCACTCGCAC CCAGAGCGCC
      ATAGTAACGT CGTGACCCCG

2591  CAGATGGTAA GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG
      GCAACTATGG ATGAACGAAA
          GTCTACCATT CGGGAGGGCA TAGCATCAAT AGATGTGCTG CCCCTCAGTC
      CGTTGATACC TACTTGCTTT

2661  TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT
      CAGACCAAGT TTACTCATAT
          ATCTGTCTAG CGACTCTATC CACGGAGTGA CTAATTCGTA ACCATTGACA
      GTCTGGTTCA AATGAGTATA
```

Fig. 45D

2731　ATACTTTAGA TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT
GAAGATCCTT TTTGATAATC
　　　　　TATGAAATCT AACTAAATTT TGAAGTAAAA ATTAAATTTT CCTAGATCCA
CTTCTAGGAA AAACTATTAG

2801　TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG AGCGTCAGAC
CCCGTAGAAA AGATCAAAGG
　　　　　AGTACTGGTT TTAGGGAATT GCACTCAAAA GCAAGGTGAC TCGCAGTCTG
GGGCATCTTT TCTAGTTTCC

2871　ATCTTCTTGA GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAAACAA
AAAAACCACC GCTACCAGCG
　　　　　TAGAAGAACT CTAGGAAAAA AGACGCGCA TTAGACGACG AACGTTTGTT
TTTTTGGTGG CGATGGTCGC

2941　GTGGTTTGTT TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAAC
TGGCTTCAGC AGAGCGCAGA
　　　　　CACCAAACAA ACGGCCTAGT TCTCGATGGT TGAGAAAAAG GCTTCCATTG
ACCGAAGTCG TCTCGCGTCT

3011　TACCAAATAC TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG
AACTCTGTAG CACCGCCTAC
　　　　　ATGGTTTATG ACAGGAAGAT CACATCGGCA TCAATCCGGT GGTGAAGTTC
TTGAGACATC GTGGCGGATG

3081　ATACCTCGCT CTGCTAATCC TGTTACCAGT GGCTGCTGCC AGTGGCGATA
AGTCGTGTCT TACCGGGTTG
　　　　　TATGGAGCGA GACGATTAGG ACAATGGTCA CCGACGACGG TCACCGCTAT
TCAGCACAGA ATGGCCCAAC

3151　GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAACGGG
GGGTTCGTGC ACACAGCCCA
　　　　　CTGAGTTCTG CTATCAATGG CCTATTCCGC GTCGCCAGCC CGACTTGCCC
CCCAAGCACG TGTGTCGGGT

3221　GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA GCGTGAGCTA
TGAGAAAGCG CCACGCTTCC
　　　　　CGAACCTCGC TTGCTGGATG TGGCTTGACT CTATGGATGT CGCACTCGAT
ACTCTTTCGC GGTGCGAAGG

3291　CGAAGGGAGA AAGGCGGACA GGTATCCGGT AAGCGGCAGG GTCGGAACAG
GAGAGCGCAC GAGGGAGCTT
　　　　　GCTTCCCTCT TTCCGCCTGT CCATAGGCCA TTCGCCGTCC AGCCTTGTC
CTCTCGCGTG CTCCCTCGAA

3361　CCAGGGGGAA ACGCCTGGTA TCTTTATAGT CCTGTCGGGT TTCGCCACCT
CTGACTTGAG CGTCGATTTT
　　　　　GGTCCCCCTT TGCGGACCAT AGAAATATCA GGACAGCCCA AAGCGGTGGA
GACTGAACTC GCAGCTAAAA

3431　TGTGATGCTC GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG
GCCTTTTTAC GGTTCCTGGC

Fig. 45E

```
        ACACTACGAG CAGTCCCCCC GCCTCGGATA CCTTTTTGCG GTCGTTGCGC
CGGAAAAATG CCAAGGACCG

   3501 CTTTTGCTGG CCTTTTGCTC ACATGTTCTT TCCTGCGTTA TCCCCTGATT
CTGTGGATAA CCGTATTACC
        GAAAACGACC GGAAAACGAG TGTACAAGAA AGGACGCAAT AGGGGACTAA
GACACCTATT GGCATAATGG

   3571 GCCTTTGAGT GAGCTGATAC CGCTCGCCGC AGCCGAACGA CCGAGCGCAG
CGAGTCAGTG AGCGAGGAAG
        CGGAAACTCA CTCGACTATG GCGAGCGGCG TCGGCTTGCT GGCTCGCGTC
GCTCAGTCAC TCGCTCCTTC

   3641 CGGAAGAGCG CCCAATACGC AAACCGCCTC TCCCCGCGCG TTGGCCGATT
CATTAATGCA GCTGGCACGA
        GCCTTCTCGC GGGTTATGCG TTTGGCGGAG AGGGGCGCGC AACCGGCTAA
GTAATTACGT CGACCGTGCT

   3711 CAGGTTTCCC GACTGGAAAG CGGGCAGTGA GCGCAACGCA ATTAATGTGA
GTTAGCTCAC TCATTAGGCA
        GTCCAAAGGG CTGACCTTTC GCCCGTCACT CGCGTTGCGT TAATTACACT
CAATCGAGTG AGTAATCCGT

   3781 CCCCAGGCTT TACACTTTAT GCTTCCGGCT CGTATGTTGT GTGGAATTGT
GAGCGGATAA CAATTTCACA
        GGGGTCCGAA ATGTGAAATA CGAAGGCCGA GCATACAACA CACCTTAACA
CTCGCCTATT GTTAAAGTGT


                                          BssHII

Kpnl                                      ~~~~~~

~~
   3851 CAGGAAACAG CTATGACCAT GATTACGCCA AGCGCGCAAT TAACCCTCAC
TAAAGGGAAC AAAAGCTGGG
        GTCCTTTGTC GATACTGGTA CTAATGCGGT TCGCGCGTTA ATTGGGAGTG
ATTTCCCTTG TTTTCGACCC


        Kpnl            XhoI
        ~~~~            ~~~~~~
   3921 TACCGGGCCC CCCCTCGAGG TCATTCATAT GCTTGAGAAG AGAGTCGGGA
TAGTCCAAAA TAAAACAAAG
        ATGGCCCGGG GGGGAGCTCC AGTAAGTATA CGAACTCTTC TCTCAGCCCT
ATCAGGTTTT ATTTTGTTTC

   3991 GTAAGATTAC CTGGTCAAAA GTGAAAACAT CAGTTAAAAG GTGGTATAAG
TAAAATATCG GTAATAAAAG
        CATTCTAATG GACCAGTTTT CACTTTTGTA GTCAATTTTC CACCATATTC
ATTTTATAGC CATTATTTTC

   4061 GTGGCCCAAA GTGAAATTTA CTCTTTTCTA CTATTATAAA AATTGAGGAT
GTTTTGTCGG TACTTTGATA
        CACCGGGTTT CACTTTAAAT GAGAAAAGAT GATAATATTT TTAACTCCTA
CAAAACAGCC ATGAAACTAT
```

Fig. 45F

4131 CGTCATTTTT GTATGAATTG GTTTTTAAGT TTATTCGCGA TTTGGAAATG
CATATCTGTA TTTGAGTCGG
        GCAGTAAAAA CATACTTAAC CAAAAATTCA AATAAGCGCT AAACCTTTAC
GTATAGACAT AAACTCAGCC

4201 TTTTTAAGTT CGTTGCTTTT GTAAATACAG AGGGATTTGT ATAAGAAATA
TCTTTAAAAA ACCCATATGC
        AAAAATTCAA GCAACGAAAA CATTTATGTC TCCCTAAACA TATTCTTTAT
AGAAATTTTT TGGGTATACG

                                                    EcoRI
                                                    ~~~~~~~
4271 TAATTTGACA TAATTTTTGA GAAAAATATA TATTCAGGCG AATTCCACAA
TGAACAATAA TAAGATTAAA
        ATTAAACTGT ATTAAAAACT CTTTTTATAT ATAAGTCCGC TTAAGGTGTT
ACTTGTTATT ATTCTAATTT

4341 ATAGCTTGCC CCCGTTGCAG CGATGGGTAT TTTTTCTAGT AAAATAAAAG
ATAAACTTAG ACTCAAAACA
        TATCGAACGG GGGCAACGTC GCTACCCATA AAAAGATCA TTTTATTTTC
TATTTGAATC TGAGTTTTGT

4411 TTTACAAAAA CAACCCCTAA AGTCCTAAAG CCCAAAGTGC TATGCACGAT
CCATAGCAAG CCCAGCCCAA
        AAATGTTTTT GTTGGGGATT TCAGGATTTC GGGTTTCACG ATACGTGCTA
GGTATCGTTC GGGTCGGGTT

4481 CCCAACCCAA CCCAACCCAC CCCAGTGCAG CCAACTGGCA AATAGTCTCC
ACCCCGGCA CTATCACCGT
        GGGTTGGGTT GGGTTGGGTG GGGTCACGTC GGTTGACCGT TTATCAGAGG
TGGGGGCCGT GATAGTGGCA

4551 GAGTTGTCCG CACCACCGCA CGTCTCGCAG CCAAAAAAAA AAAAGAAAG
AAAAAAAGA AAAGAAAAA
        CTCAACAGGC GTGGTGGCGT GCAGAGCGTC GGTTTTTTTT TTTTTCTTTC
TTTTTTTTCT TTTTCTTTTT

4621 CAGCAGGTGG GTCCGGGTCG TGGGGGCCGG AAAAGCGAGG AGGATCGCGA
GCAGCGACGA GGCCCGGCCC
        GTCGTCCACC CAGGCCCAGC ACCCCCGGCC TTTTCGCTCC TCCTAGCGCT
CGTCGCTGCT CCGGGCCGGG

4691 TCCCTCCGCT TCCAAAGAAA CGCCCCCCAT CGCCACTATA TACATACCCC
CCCCTCTCCT CCCATCCCCC
        AGGGAGGCGA AGGTTTCTTT GCGGGGGGTA GCGGTGATAT ATGTATGGGG
GGGGAGAGGA GGGTAGGGGG

4761 CAACCCTACC ACCACCACCA CCACCACCTC CTCCCCCCTC GCTGCCGGAC
GACGAGCTCC TCCCCCCTCC
        GTTGGGATGG TGGTGGTGGT GGTGGTGGAG GAGGGGGGAG CGACGGCCTG
CTGCTCGAGG AGGGGGGAGG

4831 CCCTCCGCCG CCGCCGGTAA CCACCCCGCC CCTCTCCTCT TTCTTTCTCC
GTTTTTTTTT TCGTCTCGGT

Fig. 45 G

```
          GGGAGGCGGC GGCGGCCATT GGTGGGGCGG GGAGAGGAGA AAGAAAGAGG
CAAAAAAAAA AGCAGAGCCA

    4901  CTCGATCTTT GGCCTTGGTA GTTTGGGTGG GCGAGAGCGG CTTCGTCGCC
CAGATCGGTG CGCGGGAGGG
          GAGCTAGAAA CCGGAACCAT CAAACCCACC CGCTCTCGCC GAAGCAGCGG
GTCTAGCCAC GCGCCCTCCC

                                                           BamHI
                                                           ~~~~~~
    4971  GCGGGATCTC GCGGCTGGCG TCTCCGGGCG TGAGTCGGCC CGGATCCTCG
CGGGGAATGG GGCTCTCGGA
          CGCCCTAGAG CGCCGACCGC AGAGGCCCGC ACTCAGCCGG GCCTAGGAGC
GCCCCTTACC CCGAGAGCCT

          BglII
          ~~~~~~
    5041  TGTAGATCTT CTTTCTTTCT TCTTTTTGTG GTAGAATTTG AATCCCTCAG
CATTGTTCAT CGGTAGTTTT
          ACATCTAGAA GAAAGAAAGA AGAAAAACAC CATCTTAAAC TTAGGGAGTC
GTAACAAGTA GCCATCAAAA

    5111  TCTTTTCATG ATTTGTGACA AATGCAGCCT CGTGCGGAGC TTTTTTGTAG GTAG
          AGAAAAGTAC TAAACACTGT TTACGTCGGA GCACGCCTCG AAAAAACATC CATC
```

Fig. 45H

Bss HII (4929)   Eco RI (1)

ori

pJQ3
4965 bp

amp

f1 ori

Bam HI (703)
Bgl II (735)
Hin dIII (856)
RST (GOLGI)
Not I (998)
Sal I (1098)
Acc I (1099)
Bam HI (1154)
SIG-VAC-FAE-F/S
Xho I (1536)
Nco I (1623)
Sph I (1775)
Not I (1790)
Bss HII (2030)
Bss HII (2054)
Cla I (2081)
Hin dIII (2086)
Xba I (2091)
Bss HII (2141)
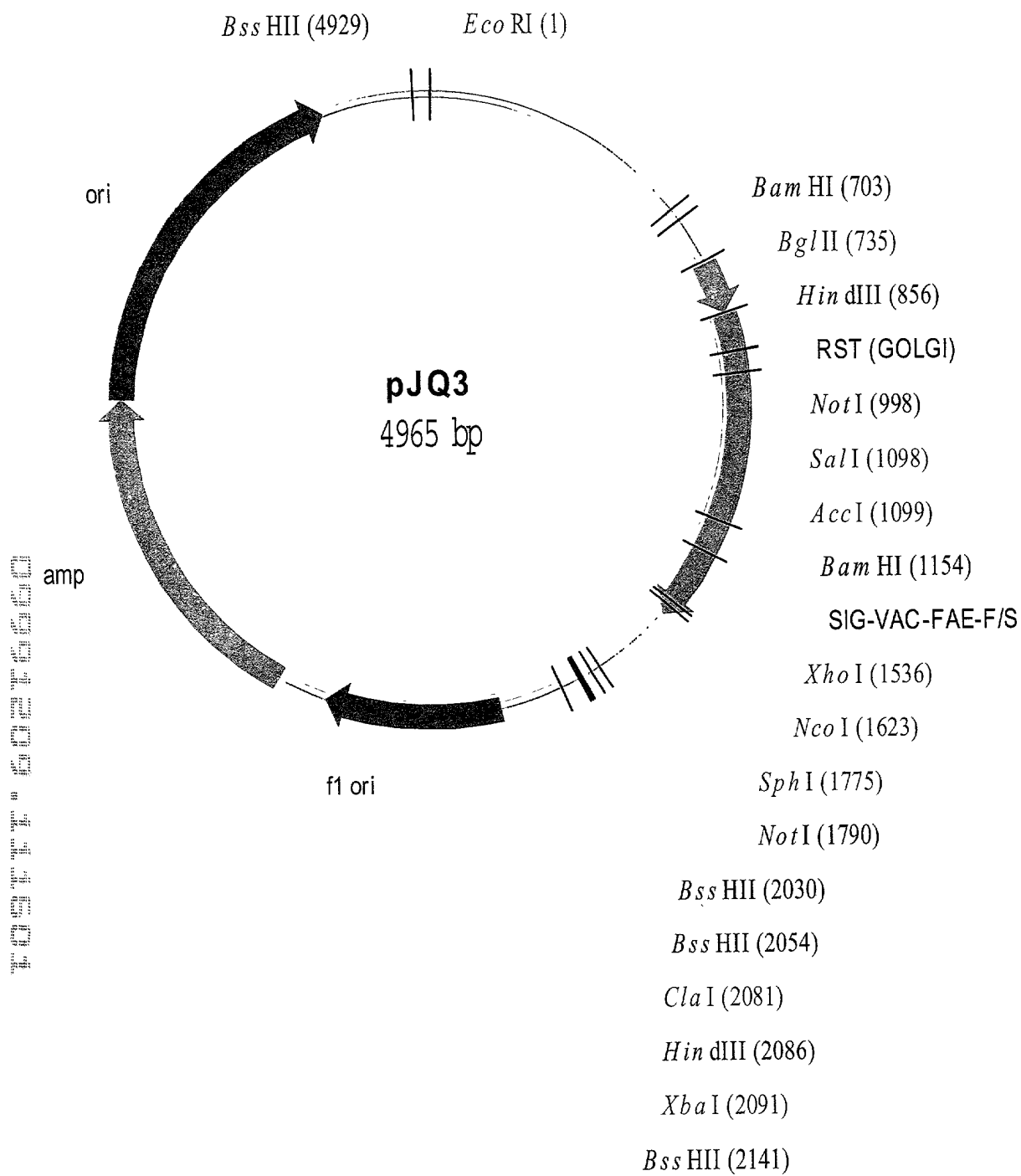
Fig. 4GA

Sequence for pJQ3

```
          EcoRI
          ~~~~~
     1    AATTCCACAA TGAACAATAA TAAGATTAAA ATAGCTTGCC CCCGTTGCAG
CGATGGGTAT TTTTTCTAGT
          TTAAGGTGTT ACTTGTTATT ATTCTAATTT TATCGAACGG GGGCAACGTC
GCTACCCATA AAAAAGATCA

    71    AAAATAAAAG ATAAACTTAG ACTCAAAACA TTTACAAAAA CAACCCCTAA
AGTCCTAAAG CCCAAAGTGC
          TTTTATTTTC TATTTGAATC TGAGTTTTGT AAATGTTTTT GTTGGGGATT
TCAGGATTTC GGGTTTCACG

   141    TATGCACGAT CCATAGCAAG CCCAGCCCAA CCCAACCCAA CCCAACCCAC
CCCAGTGCAG CCAACTGGCA
          ATACGTGCTA GGTATCGTTC GGGTCGGGTT GGGTTGGGTT GGGTTGGGTG
GGGTCACGTC GGTTGACCGT

   211    AATAGTCTCC ACCCCCGGCA CTATCACCGT GAGTTGTCCG CACCACCGCA
CGTCTCGCAG CCAAAAAAAA
          TTATCAGAGG TGGGGGCCGT GATAGTGGCA CTCAACAGGC GTGGTGGCGT
GCAGAGCGTC GGTTTTTTTT

   281    AAAAAGAAAG AAAAAAAAGA AAAAGAAAAA CAGCAGGTGG GTCCGGGTCG
TGGGGGCCGG AAAAGCGAGG
          TTTTTCTTTC TTTTTTTTCT TTTTCTTTTT GTCGTCCACC CAGGCCCAGC
ACCCCCGGCC TTTTCGCTCC

   351    AGGATCGCGA GCAGCGACGA GGCCCGGCCC TCCCTCCGCT TCCAAAGAAA
CGCCCCCCAT CGCCACTATA
          TCCTAGCGCT CGTCGCTGCT CCGGGCCGGG AGGGAGGCGA AGGTTTCTTT
GCGGGGGGTA GCGGTGATAT

   421    TACATACCCC CCCCTCTCCT CCCATCCCCC CAACCCTACC ACCACCACCA
CCACCACCTC CTCCCCCCTC
          ATGTATGGGG GGGGAGAGGA GGGTAGGGGG GTTGGGATGG TGGTGGTGGT
GGTGGTGGAG GAGGGGGGAG

   491    GCTGCCGGAC GACGAGCTCC TCCCCCCTCC CCCTCCGCCG CCGCCGGTAA
CCACCCCGCC CCTCTCCTCT
          CGACGGCCTG CTGCTCGAGG AGGGGGGAGG GGGAGGCGGC GGCGGCCATT
GGTGGGGCGG GGAGAGGAGA

   561    TTCTTTCTCC GTTTTTTTTT TCGTCTCGGT CTCGATCTTT GGCCTTGGTA
GTTTGGGTGG GCGAGAGCGG
          AAGAAAGAGG CAAAAAAAAA AGCAGAGCCA GAGCTAGAAA CCGGAACCAT
CAAACCCACC CGCTCTCGCC

   631    CTTCGTCGCC CAGATCGGTG CGCGGGAGGG GCGGGATCTC GCGGCTGGCG
TCTCCGGGCG TGAGTCGGCC
```

Fig. 46 B

```
          GAAGCAGCGG GTCTAGCCAC GCGCCCTCCC CGCCCTAGAG CGCCGACCGC
     AGAGGCCCGC ACTCAGCCGG


               BamHI                             BglII
               ~~~~~~                            ~~~~~~
      701  CGGATCCTCG CGGGGAATGG GGCTCTCGGA TGTAGATCTT CTTTCTTTCT
     TCTTTTTGTG GTAGAATTTG
               GCCTAGGAGC GCCCCTTACC CCGAGAGCCT ACATCTAGAA GAAAGAAAGA
     AGAAAAACAC CATCTTAAAC


      771  AATCCCTCAG CATTGTTCAT CGGTAGTTTT TCTTTTCATG ATTTGTGACA
     AATGCAGCCT CGTGCGGAGC
               TTAGGGAGTC GTAACAAGTA GCCATCAAAA AGAAAAGTAC TAAACACTGT
     TTACGTCGGA GCACGCCTCG


                          HindIII
                          ~~~~~~
      841  TTTTTTGTAG GTAGAAGCTT ACCATGATCC ACACCAACCT CAAAAAGAAG
     TTCTCCCTCT TCATCCTCGT
               AAAAAACATC CATCTTCGAA TGGTACTAGG TGTGGTTGGA GTTTTTCTTC
     AAGAGGGAGA AGTAGGAGCA


      911  CTTCCTCCTC TTCGCCGTGA TCTGCGTGTG GAAGAAGGGC TCCGACTACG
     AGGCCCTCAC CCTCCAAGCC
               GAAGGAGGAG AAGCGGCACT AGACGCACAC CTTCTTCCCG AGGCTGATGC
     TCCGGGAGTG GGAGGTTCGG


                          NotI
                          ~~~~~~~~~
      981  AAGGAGTTCC AAATGGCGGC CGCCTCCACG CAGGGCATCT CCGAAGACCT
     CTACAGCCGT TTAGTCGAAA
               TTCCTCAAGG TTTACCGCCG GCGGAGGTGC GTCCCGTAGA GGCTTCTGGA
     GATGTCGGCA AATCAGCTTT


                                                            SalI
                                                            ~~~~~~~
                                                            AccI
                                                            ~~~~~~~
     1051  TGGCCACTAT CTCCCAAGCT GCCTACGCCG ACCTGTGCAA CATTCCGTCG
     ACTATTATCA AGGGAGAGAA
               ACCGGTGATA GAGGGTTCGA CGGATGCGGC TGGACACGTT GTAAGGCAGC
     TGATAATAGT TCCCTCTCTT


                                                  BamHI
                                                  ~~~~~~
     1121  AATTTACAAT TCTCAAACTG ACATTAACGG ATGGATCCTC CGCGACGACA
     GCAGCAAAGA AATAATCACC
               TTAAATGTTA AGAGTTTGAC TGTAATTGCC TACCTAGGAG GCGCTGCTGT
     CGTCGTTTCT TTATTAGTGG


     1191  GTCTTCCGTG GCACTGGTAG TGATACGAAT CTACAACTCG ATACTAACTA
     CACCCTCACG CCTTTCGACA
               CAGAAGGCAC CGTGACCATC ACTATGCTTA GATGTTGAGC TATGATTGAT
     GTGGGAGTGC GGAAAGCTGT
```

Fig. 46 C

1261   CCCTACCACA ATGCAACGGT TGTGAAGTAC ACGGTGGATA TTATATTGGA
TGGGTCTCCG TCCAGGACCA
           GGGATGGTGT TACGTTGCCA ACACTTCATG TGCCACCTAT AATATAACCT
ACCCAGAGGC AGGTCCTGGT

     1331   AGTCGAGTCG CTTGTCAAAC AGCAGGTTAG CCAGTATCCG GACTACGCGC
TGACCGTGAC CGGCCACKCC
           TCAGCTCAGC GAACAGTTTG TCGTCCAATC GGTCATAGGC CTGATGCGCG
ACTGGCACTG GCCGGTGMGG

     1401   CTCGGCGCCT CCCTGGCGGC ACTCACTGCC GCCCAGCTGT CTGCGACATA
CGACAACATC CGCCTGTACA
           GAGCCGCGGA GGGACCGCCG TGAGTGACGG CGGGTCGACA GACGCTGTAT
GCTGTTGTAG GCGGACATGT

XhoI

~~~~~~
     1471   CCTTCGGCGA ACCGCGCAGC GGCAATCAGG CCTTCGCGTC GTACATGAAC
GATGCCTTCC AAGCCTCGAG
           GGAAGCCGCT TGGCGCGTCG CCGTTAGTCC GGAAGCGCAG CATGTACTTG
CTACGGAAGG TTCGGAGCTC

     1541   CCCAGATACG ACGCAGTATT CCGGGTCAC TCATGCCAAC GACGGCATCC
CAAACCTGCC CCCGGTGGAG
           GGGTCTATGC TGCGTCATAA AGGCCCAGTG AGTACGGTTG CTGCCGTAGG
GTTTGGACGG GGGCCACCTC

                        NcoI
                        ~~~~~~
     1611   CAGGGGTACG CCCATGGCGG TGTAGAGTAC TGGAGCGTTG ATCCTTACAG
CGCCCAGAAC ACATTTGTCT
           GTCCCCATGC GGGTACCGCC ACATCTCATG ACCTCGCAAC TAGGAATGTC
GCGGGTCTTG TGTAAACAGA

     1681   GCACTGGGGA TGAAGTGCAG TGCTGTGAGG CCCAGGGCGG ACAGGGTGTG
AATAATGCGC ACACGACTTA
           CGTGACCCCT ACTTCACGTC ACGACACTCC GGGTCCCGCC TGTCCCACAC
TTATTACGCG TGTGCTGAAT

                        SphI   NotI
                        ~~~~~~~   ~~~~~~~~~
     1751   TTTTGGGATG ACGAGCGGCG CATGCACCTG GCCGGTCGCG GCCGCGGAAA
CCACTGAAGG ATGAGCTGTA
           AAAACCCTAC TGCTCGCCGC GTACGTGGAC CGGCCAGCGC CGGCGCCTTT
GGTGACTTCC TACTCGACAT

     1821   AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT AAAGTTTCTT AAGATTGAAT
CCTGTTGCCG GTCTTGCGAT
           TTCTTCGTCT AGCAAGTTTG TAAACCGTTA TTTCAAAGAA TTCTAACTTA
GGACAACGGC CAGAACGCTA

                           Fig. 46D

1891 GATTATCATA TAATTTCTGT TGAATTACGT TAAGCATGTA ATAATTAACA
TGTAATGCAT GACGTTATTT
          CTAATAGTAT ATTAAAGACA ACTTAATGCA ATTCGTACAT TATTAATTGT
ACATTACGTA CTGCAATAAA

BssHII

~~
    1961 ATGAGATGGG TTTTTATGAT TAGAGTCCCG CAATTATACA TTTAATACGC
GATAGAAAAC AAAATATAGC
          TACTCTACCC AAAAATACTA ATCTCAGGGC GTTAATATGT AAATTATGCG
CTATCTTTTG TTTTATATCG

XbaI

~~~~~~~
          BssHII   BssHII   ClaI
HindIII
          ~~~~   ~~~~~~
~~~~~~~~~~~~~~
    2031 GCGCAAACTA GGATAAATTA TCGCGCGCGG TGTCATCTAT GTTACTAGAT
CGATAAGCTT CTAGAGCGGC
          CGCGTTTGAT CCTATTTAAT AGCGCGCGCC ACAGTAGATA CAATGATCTA
GCTATTCGAA GATCTCGCCG

                                                        BssHII
                                                        ~~~~~~~
    2101 CGGTGGAGCT CCAATTCGCC CTATAGTGAG TCGTATTACG CGCGCTCACT
GGCCGTCGTT TTACAACGTC
          GCCACCTCGA GGTTAAGCGG GATATCACTC AGCATAATGC GCGCGAGTGA
CCGGCAGCAA AATGTTGCAG

    2171 GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT
CCCCCTTTCG CCAGCTGGCG
          CACTGACCCT TTTGGGACCG CAATGGGTTG AATTAGCGGA ACGTCGTGTA
GGGGGAAAGC GGTCGACCGC

    2241 TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC
TGAATGGCGA ATGGGACGCG
          ATTATCGCTT CTCCGGGCGT GGCTAGCGGG AAGGGTTGTC AACGCGTCGG
ACTTACCGCT TACCCTGCGC

    2311 CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT
GACCGCTACA CTTGCCAGCG
          GGGACATCGC CGCGTAATTC GCGCCGCCCA CACCACCAAT GCGCGTCGCA
CTGGCGATGT GAACGGTCGC

    2381 CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT CGCCACGTTC
GCCGGCTTTC CCCGTCAAGC
          GGGATCGCGG GCGAGGAAAG CGAAAGAAGG GAAGGAAAGA GCGGTGCAAG
CGGCCGAAAG GGGCAGTTCG

Fig. 46E

```
2451   TCTAAATCGG GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC
TCGACCCCAA AAAACTTGAT
       AGATTTAGCC CCCGAGGGAA ATCCCAAGGC TAAATCACGA AATGCCGTGG
AGCTGGGGTT TTTTGAACTA

2521   TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG
CCCTTTGACG TTGGAGTCCA
       ATCCCACTAC CAAGTGCATC ACCCGGTAGC GGGACTATCT GCCAAAAAGC
GGGAAACTGC AACCTCAGGT

2591   CGTTCTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCT
ATCTCGGTCT ATTCTTTTGA
       GCAAGAAATT ATCACCTGAG AACAAGGTTT GACCTTGTTG TGAGTTGGGA
TAGAGCCAGA TAAGAAAACT

2661   TTTATAAGGG ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA
TTTAACAAAA ATTTAACGCG
       AAATATTCCC TAAAACGGCT AAAGCCGGAT AACCAATTTT TTACTCGACT
AAATTGTTTT TAAATTGCGC

2731   AATTTTAACA AAATATTAAC GCTTACAATT TAGGTGGCAC TTTTCGGGGA
AATGTGCGCG GAACCCCTAT
       TTAAAATTGT TTTATAATTG CGAATGTTAA ATCCACCGTG AAAAGCCCCT
TTACACGCGC CTTGGGGATA

2801   TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
AACCCTGATA AATGCTTCAA
       AACAAATAAA AAGATTTATG TAAGTTTATA CATAGGCGAG TACTCTGTTA
TTGGGACTAT TTACGAAGTT

2871   TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT
TATTCCCTTT TTTGCGGCAT
       ATTATAACTT TTTCCTTCTC ATACTCATAA GTTGTAAAGG CACAGCGGGA
ATAAGGGAAA AAACGCCGTA

2941   TTTGCCTTCC TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT
GCTGAAGATC AGTTGGGTGC
       AAACGGAAGG ACAAAAACGA GTGGGTCTTT GCGACCACTT TCATTTTCTA
CGACTTCTAG TCAACCCACG

3011   ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA
GTTTTCGCCC CGAAGAACGT
       TGCTCACCCA ATGTAGCTTG ACCTAGAGTT GTCGCCATTC TAGGAACTCT
CAAAAGCGGG GCTTCTTGCA

3081   TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC
CCGTATTGAC GCCGGGCAAG
       AAAGGTTACT ACTCGTGAAA ATTTCAAGAC GATACACCGC GCCATAATAG
GGCATAACTG CGGCCCGTTC

3151   AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC
TCACCAGTCA CAGAAAAGCA
       TCGTTGAGCC AGCGGCGTAT GTGATAAGAG TCTTACTGAA CCAACTCATG
AGTGGTCAGT GTCTTTTCGT
```

Fig. 46F

3221 TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA
TGAGTGATAA CACTGCGGCC
         AGAATGCCTA CCGTACTGTC ATTCTCTTAA TACGTCACGA CGGTATTGGT
ACTCACTATT GTGACGCCGG

3291 AACTTACTTC TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT
GCACAACATG GGGGATCATG
         TTGAATGAAG ACTGTTGCTA GCCTCCTGGC TTCCTCGATT GGCGAAAAAA
CGTGTTGTAC CCCCTAGTAC

3361 TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC
GACGAGCGTG ACACCACGAT
         ATTGAGCGGA ACTAGCAACC CTTGGCCTCG ACTTACTTCG GTATGGTTTG
CTGCTCGCAC TGTGGTGCTA

3431 GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC
TTACTCTAGC TTCCCGGCAA
         CGGACATCGT TACCGTTGTT GCAACGCGTT TGATAATTGA CCGCTTGATG
AATGAGATCG AAGGGCCGTT

3501 CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC CACTTCTGCG
CTCGGCCCTT CCGGCTGGCT
         GTTAATTATC TGACCTACCT CCGCCTATTT CAACGTCCTG GTGAAGACGC
GAGCCGGGAA GGCCGACCGA

3571 GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC
ATTGCAGCAC TGGGGCCAGA
         CCAAATAACG ACTATTTAGA CCTCGGCCAC TCGCACCCAG AGCGCCATAG
TAACGTCGTG ACCCCGGTCT

3641 TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA
CTATGGATGA ACGAAATAGA
         ACCATTCGGG AGGGCATAGC ATCAATAGAT GTGCTGCCCC TCAGTCCGTT
GATACCTACT TGCTTTATCT

3711 CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA
CCAAGTTTAC TCATATATAC
         GTCTAGCGAC TCTATCCACG GAGTGACTAA TTCGTAACCA TTGACAGTCT
GGTTCAAATG AGTATATATG

3781 TTTAGATTGA TTTAAAACTT CATTTTTAAT TTAAAAGGAT CTAGGTGAAG
ATCCTTTTTG ATAATCTCAT
         AAATCTAACT AAATTTTGAA GTAAAAATTA AATTTTCCTA GATCCACTTC
TAGGAAAAAC TATTAGAGTA

3851 GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG
TAGAAAAGAT CAAAGGATCT
         CTGGTTTTAG GGAATTGCAC TCAAAAGCAA GGTGACTCGC AGTCTGGGGC
ATCTTTTCTA GTTTCCTAGA

3921 TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA
ACCACCGCTA CCAGCGGTGG

Fig. 46 G

```
        AGAACTCTAG GAAAAAAAGA CGCGCATTAG ACGACGAACG TTTGTTTTTT
TGGTGGCGAT GGTCGCCACC

   3991  TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTTCCGAA GGTAACTGGC
TTCAGCAGAG CGCAGATACC
        AAACAAACGG CCTAGTTCTC GATGGTTGAG AAAAAGGCTT CCATTGACCG
AAGTCGTCTC GCGTCTATGG

   4061  AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT
CTGTAGCACC GCCTACATAC
        TTTATGACAG GAAGATCACA TCGGCATCAA TCCGGTGGTG AAGTTCTTGA
GACATCGTGG CGGATGTATG

   4131  CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC
GTGTCTTACC GGGTTGGACT
        GAGCGAGACG ATTAGGACAA TGGTCACCGA CGACGGTCAC CGCTATTCAG
CACAGAATGG CCCAACCTGA

   4201  CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT
TCGTGCACAC AGCCCAGCTT
        GTTCTGCTAT CAATGGCCTA TTCCGCGTCG CCAGCCCGAC TTGCCCCCCA
AGCACGTGTG TCGGGTCGAA

   4271  GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG
AAAGCGCCAC GCTTCCCGAA
        CCTCGCTTGC TGGATGTGGC TTGACTCTAT GGATGTCGCA CTCGATACTC
TTTCGCGGTG CGAAGGGCTT

   4341  GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG AACAGGAGA
GCGCACGAGG GAGCTTCCAG
        CCCTCTTTCC GCCTGTCCAT AGGCCATTCG CCGTCCCAGC CTTGTCCTCT
CGCGTGCTCC CTCGAAGGTC

   4411  GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA
CTTGAGCGTC GATTTTTGTG
        CCCCTTTGCG GACCATAGAA ATATCAGGAC AGCCCAAAGC GGTGGAGACT
GAACTCGCAG CTAAAAACAC

   4481  ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCCT
TTTTACGGTT CCTGGCCTTT
        TACGAGCAGT CCCCCCGCCT CGGATACCTT TTTGCGGTCG TTGCGCCGGA
AAAATGCCAA GGACCGGAAA

   4551  TGCTGGCCTT TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT
GGATAACCGT ATTACCGCCT
        ACGACCGGAA AACGAGTGTA CAAGAAAGGA CGCAATAGGG GACTAAGACA
CCTATTGGCA TAATGGCGGA

   4621  TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG
TCAGTGAGCG AGGAAGCGGA
        AACTCACTCG ACTATGGCGA GCGGCGTCGG CTTGCTGGCT CGCGTCGCTC
AGTCACTCGC TCCTTCGCCT
```

Fig. 46 H

4691 AGAGCGCCCA ATACGCAAAC CGCCTCTCCC CGCGCGTTGG CCGATTCATT
AATGCAGCTG GCACGACAGG
         TCTCGCGGGT TATGCGTTTG GCGGAGAGGG GCGCGCAACC GGCTAAGTAA
TTACGTCGAC CGTGCTGTCC

4761 TTTCCCGACT GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA
GCTCACTCAT TAGGCACCCC
         AAAGGGCTGA CCTTTCGCCC GTCACTCGCG TTGCGTTAAT TACACTCAAT
CGAGTGAGTA ATCCGTGGGG

4831 AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTTGTGTGG AATTGTGAGC
GGATAACAAT TTCACACAGG
         TCCGAAATGT GAAATACGAA GGCCGAGCAT ACAACACACC TTAACACTCG
CCTATTGTTA AAGTGTGTCC

                                              BssHII

EcoRI
                                              ~~~~~~~
4901 AAACAGCTAT GACCATGATT ACGCCAAGCG CGCAATTAAC CCTCACTAAA
GGGAACAAAA GCTGG
         TTTGTCGATA CTGGTACTAA TGCGGTTCGC GCGTTAATTG GGAGTGATTT
CCCTTGTTTT CGACC

Fig. 4C I

Figure 47 A

Hin dIII (2)

Nco I (9)

aleurain signal - NPIR deletion

Bgl II (5179)

Not I (87)

Sal I (187)

Rice actin

Bam HI (5147)

Bam HI (243)

FAE

Eco RI (4445)

Xho I (625)

Nco I (712)

Xho I (4069)

pUG4

Sph I (864)

Kpn I (4058)

5295 bp

Not I (879)

nos   KDEL

Cla I (1169)

Hin dIII (1174)

ori

Xba I (1179)

f1 ori

amp
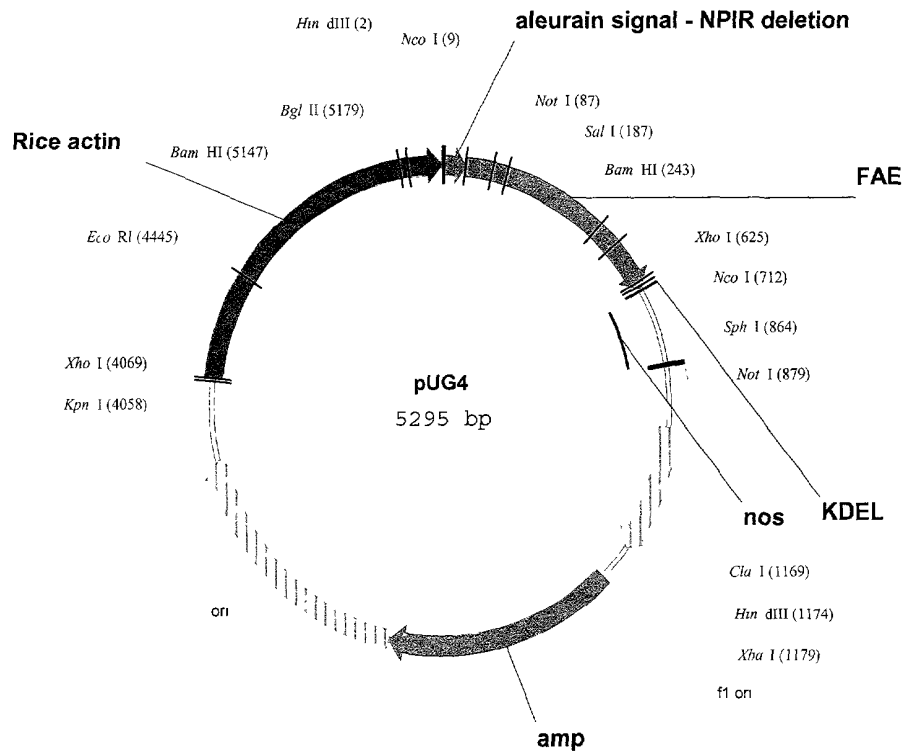
Figure 47B

```
                    NcoI
                   ~~~~~~~
       HindIII
       ~~~~~
             M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
  1   AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG
                     NotI
                    ~~~~~~~~
             · A   S   S   R   A   A   A   S   T   Q   G   I   S   E   D   L   Y   S   R   L   V   E   M ·
 71   TCGCCTCCTC CCGCGCGGCC GCCTCCACGC AGGGCATCTC CGAAGACCTC TACAGCCGTT TAGTCGAAAT
                                                                         SalI
                                                                        ~~~~~~~
             · A   T   I   S   Q   A   A   Y   A   D   L   C   N   I   P   S   T   I   I   K   G   E   K
141   GGCCACTATC TCCCAAGCTG CCTACGCCGA CCTGTGCAAC ATTCCGTCGA CTATTATCAA GGGAGAGAAA
                                                  BamHI
                                                 ~~~~~~~
             I   Y   N   S   Q   T   D   I   N   G   W   I   L   R   D   D   S   K   E   I   I   T   V
211   ATTTACAATT CTCAAACTGA CATTAACGGA TGGATCCTCC GCGACGACAG CAGCAAAGAA ATAATCACCG
             · F   R   G   T   G   S   D   T   N   L   Q   L   D   T   N   Y   T   L   T   P   F   D   T ·
281   TCTTCCGTGG CACTGGTAGT GATACGAATC TACAACTCGA TACTAACTAC ACCCTCACGC CTTTCGACAC
             · L   P   Q   C   N   G   C   E   V   H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q
351   CCTACCACAA TGCAACGGTT GTGAAGTACA CGGTGGATAT TATATTGGAT GGGTCTCCGT CCAGGACCAA
             V   E   S   L   V   K   Q   Q   V   S   Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L
421   GTCGAGTCGC TTGTCAAACA GCAGGTTAGC CAGTATCCGG ACTACGCGCT GACCGTGACC GGCCACKCCC
             · G   A   S   L   A   A   L   T   A   A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T ·
491   TCGGCGCCTC CCTGGCGGCA CTCACTGCCG CCCAGCTGTC TGCGACATAC GACAACATCC GCCTGTACAC
                                                                                     XhoI
                                                                                      ~
             · F   G   E   P   R   S   G   N   Q   A   F   A   S   Y   M   N   D   A   F   Q   A   S   S
561   CTTCGGCGAA CCGCGCAGCG GCAATCAGGC CTTCGCGTCG TACATGAACG ATGCCTTCCA AGCCTCGAGC
             P   D   T   T   Q   Y   F   R   V   T   H   A   N   D   G   I   P   N   L   P   P   V   E   Q
631   CCAGATACGA CGCAGTATTT CCGGGTCACT CATGCCAACG ACGGCATCCC AAACCTGCCC CCGGTGGAGC
                     NcoI
                    ~~~~~~
             · G   Y   A   H   G   G   V   E   Y   W   S   V   D   P   Y   S   A   Q   N   T   F   V   C ·
701   AGGGGTACGC CCATGGCGGT GTAGAGTACT GGAGCGTTGA TCCTTACAGC GCCCAGAACA CATTTGTCTG
             · T   G   D   E   V   Q   C   C   E   A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y
771   CACTGGGGAT GAAGTGCAGT GCTGTGAGGC CCAGGGCGGA CAGGGTGTGA ATAATGCGCA CACGACTTAT
                         SphI                       NotI
                        ~~~~~~                    ~~~~~~~~
             F   G   M   T   S   G   A   C   T   W   P   V   A   A   A   E   P   L   K   D   E   L   *
841   TTTGGGATGA CGAGCGGCGC ATGCACCTGG CCGGTCGCGG CCGCGGAACC ACTGAAGGAT GAGCTGTAAA
911   GAAGCAGATC GTTCAAACAT TTGGCAATAA AGTTTCTTAA GATTGAATCC TGTTGCCGGT CTTGCGATGA
981   TTATCATATA ATTTCTGTTG AATTACGTTA AGCATGTAAT AATTAACATG TAATGCATGA CGTTATTTAT
1051  GAGATGGGTT TTTATGATTA GAGTCCCGCA ATTATACATT TAATACGCGA TAGAAACAA AAATATAGCGC
                                                              HindIII
                                                             ~~~~~~~
                                                     ClaI          XbaI
                                                     ~~~~~~        ~~~~~~
1121  GCAAACTAGG ATAAATTATC GCGCGCGGTG TCATCTATGT TACTAGATCG ATAAGCTTCT AGAGCGGCCG
1191  GTGGAGCTCC AATTCGCCCT ATAGTGAGTC GTATTACGCG CGCTCACTGG CCGTCGTTTT ACAACGTCGT
1261  GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC CCCTTTCGCC AGCTGGCGTA
1331  ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGGACGCGCC
1401  CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG CGCAGCGTGA CCGCTACACT TGCCAGCGCC
1471  CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT TCCTTTCTCG CCACGTTCGC CGGCTTTCCC CGTCAAGCTC
1541  TAAATCGGGG GCTCCCTTTA GGGTTCCGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTA
1611  GGGTGATGGT TCACGTAGTG GGCCATCGCC CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG
1681  TTCTTTAATA GTGGACTCTT GTTCCAAACT GGAACAACAC TCAACCCTAT CTCGGTCTAT TCTTTTGATT
1751  TATAAGGGAT TTTGCCGATT TCGGCCTATT GGTTAAAAAA TGAGCTGATT TAACAAAAAT TTAACGCGAA
```

Fig. 47 C

```
1821    TTTTAACAAA ATATTAACGC TTACAATTTA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT
1891    GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA TGCTTCAATA
1961    ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA TTCCCTTTTT TGCGGCATTT
2031    TGCCTTCCTG TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC
2101    GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTCGCCCCG AAGAACGTTT
2171    TCCAATGATG AGCACTTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG
2241    CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAAGCATC
2311    TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA
2381    CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG GGATCATGTA
2451    ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCACGATGC
2521    CTGTAGCAAT GGCAACAACG TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA
2591    ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG
2661    TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG
2731    GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
2801    GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC AAGTTTACTC ATATATACTT
2871    TAGATTGATT TAAAACTTCA TTTTTAATTT AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA
2941    CCAAATCCC TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC
3011    TTGAGATCCT TTTTTTCTGC GCGTAATCTG CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT
3081    TGTTTGCCGG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA
3151    ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC CTACATACCT
3221    CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG GTTGGACTCA
3291    AGACGATAGT TACCGGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTTC GTGCACACAG CCCAGCTTGG
3361    AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA GCTATGAGAA AGCGCCACGC TTCCCGAAGG
3431    GAGAAAGGCG GACAGGTATC CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG
3501    GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTTGTGAT
3571    GCTCGTCAGG GGGGCGGAGC CTATGGAAAA ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG
3641    CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG ATAACCGTAT TACCGCCTTT
3711    GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG
3781    AGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC ACGACAGGTT
3851    TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC TCACTCATTA GGCACCCCAG
3921    GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA
                                                                                     KpnI
                                                                                      ~~

3991    ACAGCTATGA CCATGATTAC GCCAAGCGCG CAATTAACCC TCACTAAAGG GAACAAAAGC TGGGTACCGG
                   XhoI
                  ~~~~~~~
4061    GCCCCCCCTC GAGGTCATTC ATATGCTTGA GAAGAGAGTC GGGATAGTCC AAAATAAAAC AAAGGTAAGA
4131    TTACCTGGTC AAAAGTGAAA ACATCAGTTA AAAGGTGGTA TAAGTAAAAT ATCGGTAATA AAAGGTGGCC
4201    CAAAGTGAAA TTTACTCTTT TCTACTATTA TAAAAATTGA GGATGTTTTG TCGGTACTTT GATACGTCAT
4271    TTTTGTATGA ATTGGTTTTT AAGTTTATTC GCGATTTGGA AATGCATATC TGTATTTGAG TCGGTTTTTA
4341    AGTTCGTTGC TTTTGTAAAT ACAGAGGGAT TTGTATAAGA AATATCTTTA AAAACCCAT ATGCTAATTT
                                                  EcoRI
                                                 ~~~~~~~
4411    GACATAATTT TTGAGAAAAA TATATATTCA GGCGAATTCC ACAATGAACA ATAATAAGAT TAAAATAGCT
4481    TGCCCCCGTT GCAGCGATGG GTATTTTTTC TAGTAAAATA AAAGATAAAC TTAGACTCAA AACATTTACA
4551    AAAACAACCC CTAAAGTCCT AAAGCCCAAA GTGCTATGCA CGATCCATAG CAAGCCCAGC CCAACCCAAC
4621    CCAACCCAAC CCACCCCAGT GCAGCCAACT GGCAAATAGT CTCCACCCCC GGCACTATCA CCGTGAGTTG
4691    TCCGCACCAC CGCACGTCTC GCAGCCAAAA AAAAAAAAG AAAGAAAAA AAGAAAAAGA AAAACAGCAG
4761    GTGGGTCCGG TCGTGGGGG CCGGAAAAGC GAGGAGGATC GCGAGCAGCG ACGAGGCCCG GCCCTCCCTC
4831    CGCTTCCAAA GAAACGCCCC CCATCGCCAC TATATACATA CCCCCCCCTC TCCTCCCATC CCCCCAACCC
4901    TACCACCACC ACCACCACCA CCTCCTCCCC CCTCGCTGCC GGACGACGAG CTCCTCCCCC CTCCCCCTCC
4971    GCCGCCGCCT GTAACCACCC CGCCCCTCTC CTCTTTCTTT CTCCGTTTTT TTTTTCGTCT CGGTCTCGAT
5041    CTTTGGCCTT GGTAGTTTGG GTGGGCGAGA GCGGCTTCGT CGCCCAGATC GGTGCGCGGG AGGGGCGGGA
                                                           BamHI                           BglII
                                                          ~~~~~~~
5111    TCTCGCGGCT GGCGTCTCCG GGCGTGAGTC GGCCCGGATC CTCGCGGGGA ATGGGGCTCT CGGATGTAGA
            BglII
           ~~~
5181    TCTTCTTTCT TTCTTCTTTT TGTGGTAGAA TTTGAATCCC TCAGCATTGT TCATCGGTAG TTTTTCTTTT
5251    CATGATTTGT GACAAATGCA GCCTCGTGCG GAGCTTTTTT GTAGC
```

Kpn I (1)

Nco I (1)
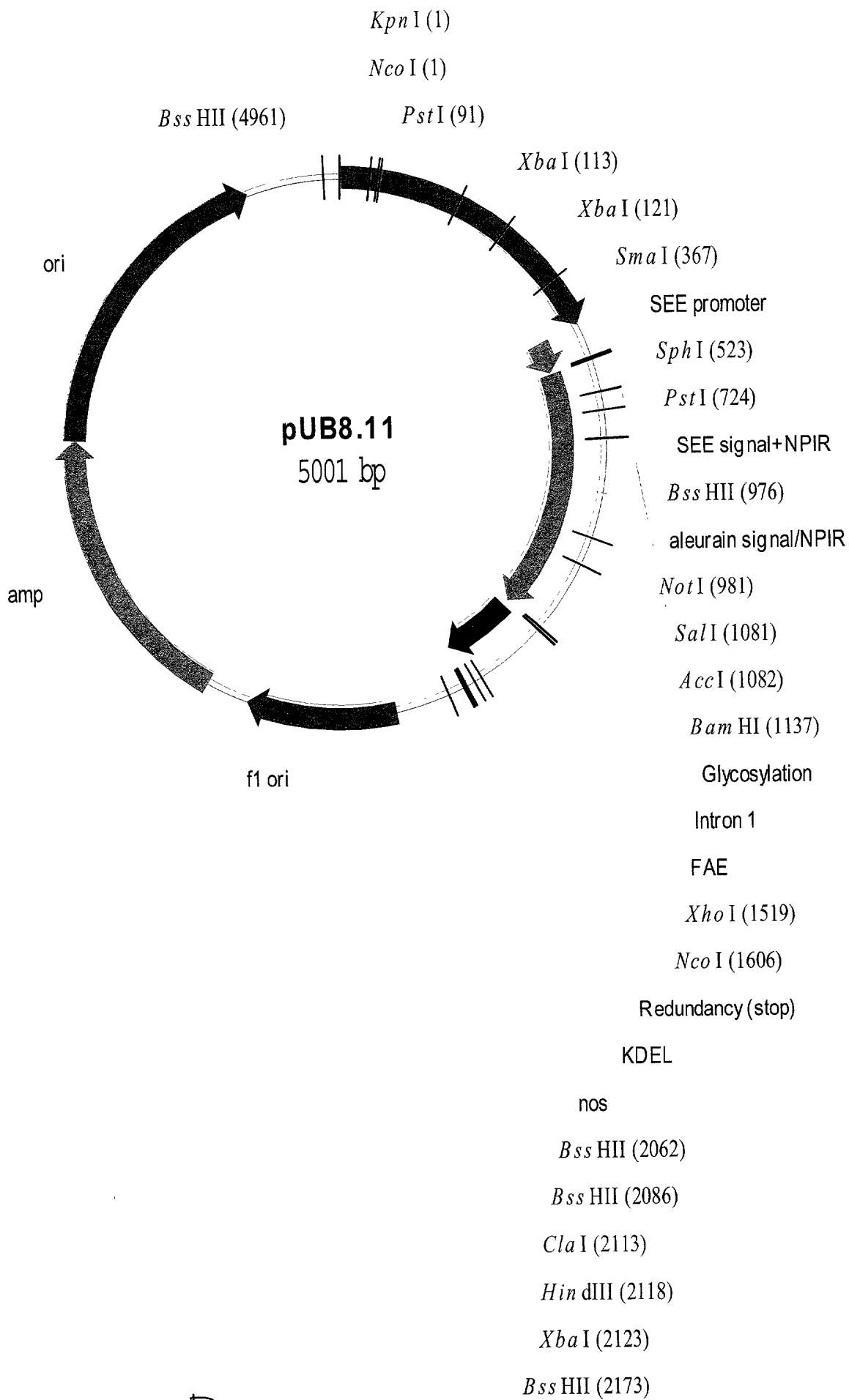
Bss HII (4961)   *Pst* I (91)

Xba I (113)

Xba I (121)

Sma I (367)

SEE promoter

Sph I (523)

Pst I (724)

SEE signal+NPIR

Bss HII (976)

aleurain signal/NPIR

Not I (981)

Sal I (1081)

Acc I (1082)

Bam HI (1137)

Glycosylation

Intron 1

FAE

Xho I (1519)

Nco I (1606)

Redundancy (stop)

KDEL

nos

Bss HII (2062)

Bss HII (2086)

Cla I (2113)

Hin dIII (2118)

Xba I (2123)

Bss HII (2173)

ori

amp

f1 ori

pUB8.11
5001 bp

Fig. 48 A

Sequence for pUB8.11

```
         NcoI
         ~~~~~
         KpnI
         ~
     1   CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC
ACCATTGGCT ACAATATCTG
         GTACCCGGTC CATATTAATA CCCTATAGAG TTCGTTTATT AGCTTTATAG
TGGTAACCGA TGTTATAGAC


                          PstI                         XbaI      XbaI
                          ~~~~~~                       ~~~~~~  ~~~~~~~
    71   AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTTGT ATCTAGAACT
CTAGATAGCA CAGCCACAGC
         TCGAGGCTCA AGACTGACGT CAGACCTACT GCGCACAACA TAGATCTTGA
GATCTATCGT GTCGGTGTCG


    141  ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTTG GAGACGGAGC
TCTTTCCTAC CTCCTGACGT
         TGGATGTCCT CACGCTGTGA ACACCTGACA TCATCACAAC CTCTGCCTCG
AGAAAGGATG GAGGACTGCA


    211  TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC
CCAACAAAAT ATCGTCCCCC
         ACGGCGGCAA CAGGTAAGGT TGCCGTAGTG AGAGTTGGTT AGTGCGCGAG
GGTTGTTTTA TAGCAGGGGG


    281  ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTCGCGCC GTTTTTGTCT
GAATCTCGCT TCCACTGGCC
         TACAGAACCG CCTCTCTCTC ATGTATGTAC GACAGCGCGG CAAAAACAGA
CTTAGAGCGA AGGTGACCGG


                          SmaI
                          ~~~~~~
    351  AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCG
TCACCCCTGG CGTCATGGGA
         TTAGTCGAGT CGAGGGCCCT CGAGTGAGTA AGTTCTAGGG TAGCAGCAGC
AGTGGGGACC GCAGTACCCT


    421  TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA
GTACTGCAAG ATAACCCAAT
         ACCTTTTCTT GGAGGCAACG AGCCTACTCA GTCGGTATAG GGGCTTGTCT
CATGACGTTC TATTGGGTTA


                                         SphI
                                         ~~~~~~~
    491  TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTTTGG
CTTAATTGAC TTTATTTTTG
         AGTCTAAGGG GGTTATCTCT TTCATATCGT ACGAAAGCCC AAAACAAACC
GAATTAACTG AAATAAAAAC
```

Fig. 48B

561 TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT
CGAGACGGAT AATAGGCTGG
      AACCTCAACT TACGACTAAA CAACACATTT TACGGGTTGG TAGACTTATA
GCTCTGCCTA TTATCCGACC

631 CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT
GGGCATTACA GCTGGAGGCT
      GATTAATTAA ATATCGTTCT AAGACATCAC GTGTAGCGTT TATAGAAAGA
CCCGTAATGT CGACCTCCGA

                              PstI
                           ~~~~~~~
701 TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG
CGATGAGATG GGTATAAAAC
      AGTAGTCGGA CTTTGTGAGA CGTCTCGGAC TTCGTTCACC ACTTCGCACC
GCTACTCTAC CCATATTTTG

771 CCCCGGCACC GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCTCGC
TCCCCCTGCC GGACGACCCA
      GGGGCCGTGG CCCTGCGCTC GAGGGCGGAT GGTCATGGTA GAGCGGAGCG
AGGGGGACGG CCTGCTGGGT

841 GTAAAATACT GTTGCCCACT CGCCGGCGAG ATGGCCCACG GCCGCATCCT
CTTCTTGGCG CTCGCCGTCT
      CATTTTATGA CAACGGGTGA GCGGCCGCTC TACCGGGTGC CGGCGTAGGA
GAAGAACCGC GAGCGGCAGA

BssHII

~~~~~~

NotI

~~
911 TGGCCACCGC CGCGGTGGCC GCCGCATCNT TGGCGGACTC CAACCCGATC
CGGCCCGTCA CCGAGCGCGC
      ACCGGTGGCG CGCCACCGG CGGCGTAGNA ACCGCCTGAG GTTGGGCTAG
GCCGGGCAGT GGCTCGCGCG

         NotI
         ~~~~~~
981 GGCCGCCTCC ACGCAGGGCA TCTCCGAAGA CCTCTACAGC CGTTTAGTCG
AAATGGCCAC TATCTCCCAA
      CCGGCGGAGG TGCGTCCCGT AGAGGCTTCT GGAGATGTCG GCAAATCAGC
TTTACCGGTG ATAGAGGGTT

                              SalI
                           ~~~~~~~
                              AccI
                           ~~~~~~~
1051 GCTGCCTACG CCGACCTGTG CAACATTCCG TCGACTATTA TCAAGGGAGA
GAAAATTTAC AATTCTCAAA

                    Fig. 48 C

CGACGGATGC GGCTGGACAC GTTGTAAGGC AGCTGATAAT AGTTCCCTCT
CTTTTAAATG TTAAGAGTTT

                        BamHI
                        ~~~~~~~
    1121   CTGACATTAA CGGATGGATC CTCCGCGACG ACAGCAGCAA AGAAATAATC
ACCGTCTTCC GTGGCACTGG
            GACTGTAATT GCCTACCTAG GAGGCGCTGC TGTCGTCGTT TCTTTATTAG
TGGCAGAAGG CACCGTGACC

    1191   TAGTGATACG AATCTACAAC TCGATACTAA CTACACCCTC ACGCCTTTCG
ACACCCTACC ACAATGCAAC
            ATCACTATGC TTAGATGTTG AGCTATGATT GATGTGGGAG TGCGGAAAGC
TGTGGGATGG TGTTACGTTG

    1261   GGTTGTGAAG TACACGGTGG ATATTATATT GGATGGGTCT CCGTCCAGGA
CCAAGTCGAG TCGCTTGTCA
            CCAACACTTC ATGTGCCACC TATAATATAA CCTACCCAGA GGCAGGTCCT
GGTTCAGCTC AGCGAACAGT

    1331   AACAGCAGGT TAGCCAGTAT CCGGACTACG CGCTGACCGT GACCGGCCAC
KCCCTCGGCG CCTCCCTGGC
            TTGTCGTCCA ATCGGTCATA GGCCTGATGC GCGACTGGCA CTGGCCGGTG
MGGGAGCCGC GGAGGGACCG

    1401   GGCACTCACT GCCGCCCAGC TGTCTGCGAC ATACGACAAC ATCCGCCTGT
ACACCTTCGG CGAACCGCGC
            CCGTGAGTGA CGGCGGGTCG ACAGACGCTG TATGCTGTTG TAGGCGGACA
TGTGGAAGCC GCTTGGCGCG

                                                        XhoI
                                                        ~~~~~~~
    1471   AGCGGCAATC AGGCCTTCGC GTCGTACATG AACGATGCCT TCCAAGCCTC
GAGCCCAGAT ACGACGCAGT
            TCGCCGTTAG TCCGGAAGCG CAGCATGTAC TTGCTACGGA AGGTTCGGAG
CTCGGGTCTA TGCTGCGTCA

NcoI

~~~~~~
    1541   ATTTCCGGGT CACTCATGCC AACGACGGCA TCCCAAACCT GCCCCCGGTG
GAGCAGGGGT ACGCCCATGG
            TAAAGGCCCA GTGAGTACGG TTGCTGCCGT AGGGTTTGGA CGGGGGCCAC
CTCGTCCCCA TGCGGGTACC

    1611   CGGTGTAGAG TACTGGAGCG TTGATCCTTA CAGCGCCCAG AACACATTTG
TCTGCACTGG GGATGAAGTG
            GCCACATCTC ATGACCTCGC AACTAGGAAT GTCGCGGGTC TTGTGTAAAC
AGACGTGACC CCTACTTCAC

    1681   CAGTGCTGTG AGGCCCAGGG CGGACAGGGT GTGAATAATG CGCACACGAC
TTATTTTGGG ATGACGAGCG

                    Fig. 48D

```
          GTCACGACAC TCCGGGTCCC GCCTGTCCCA CACTTATTAC GCGTGTGCTG
AATAAAACCC TACTGCTCGC

     1751  GAGCCTGTAC ATGGTGATCA GTCATTTCAG CCTCCCCGAG TGTACCAGGA
AAGATGGATG TCCTGGAGAG
          CTCGGACATG TACCACTAGT CAGTAAAGTC GGAGGGGCTC ACATGGTCCT
TTCTACCTAC AGGACCTCTC

     1821  GGGGCCGCGT AACCACTGAA GGATGAGCTG TAAAGAAGCA GATCGTTCAA
ACATTTGGCA ATAAAGTTTC
          CCCCGGCGCA TTGGTGACTT CCTACTCGAC ATTTCTTCGT CTAGCAAGTT
TGTAAACCGT TATTTCAAAG

     1891  TTAAGATTGA ATCCTGTTGC CGGTCTTGCG ATGATTATCA TATAATTTCT
GTTGAATTAC GTTAAGCATG
          AATTCTAACT TAGGACAACG GCCAGAACGC TACTAATAGT ATATTAAAGA
CAACTTAATG CAATTCGTAC

     1961  TAATAATTAA CATGTAATGC ATGACGTTAT TTATGAGATG GGTTTTTATG
ATTAGAGTCC CGCAATTATA
          ATTATTAATT GTACATTACG TACTGCAATA AATACTCTAC CCAAAAATAC
TAATCTCAGG GCGTTAATAT

                                                  BssHII

BssHII
                                                  ~~~~~~

~~~~~~
     2031  CATTTAATAC GCGATAGAAA ACAAAATATA GCGCGCAAAC TAGGATAAAT
TATCGCGCGC GGTGTCATCT
          GTAAATTATG CGCTATCTTT TGTTTTATAT CGCGCGTTTG ATCCTATTTA
ATAGCGCGCG CCACAGTAGA

                                       XbaI

                                       ~~~~~~
                          ClaI HindIII
                          ~~~~~~~~~~~~~~
     2101  ATGTTACTAG ATCGATAAGC TTCTAGAGCG GCCGGTGGAG CTCCAATTCG
CCCTATAGTG AGTCGTATTA
          TACAATGATC TAGCTATTCG AAGATCTCGC CGGCCACCTC GAGGTTAAGC
GGGATATCAC TCAGCATAAT

               BssHII
               ~~~~~~
     2171  CGCGCGCTCA CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCTG
GCGTTACCCA ACTTAATCGC
          GCGCGCGAGT GACCGGCAGC AAAATGTTGC AGCACTGACC CTTTTGGGAC
CGCAATGGGT TGAATTAGCG

     2241  CTTGCAGCAC ATCCCCCTTT CGCCAGCTGG CGTAATAGCG AAGAGGCCCG
CACCGATCGC CCTTCCCAAC
          GAACGTCGTG TAGGGGGAAA GCGGTCGACC GCATTATCGC TTCTCCGGGC
GTGGCTAGCG GGAAGGGTTG
```

Fig. 48 E

2311 AGTTGCGCAG CCTGAATGGC GAATGGGACG CGCCCTGTAG CGGCGCATTA
AGCGCGGCGG GTGTGGTGGT
          TCAACGCGTC GGACTTACCG CTTACCCTGC GCGGGACATC GCCGCGTAAT
TCGCGCCGCC CACACCACCA

2381 TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT
TCGCTTTCTT CCCTTCCTTT
          ATGCGCGTCG CACTGGCGAT GTGAACGGTC GCGGGATCGC GGGCGAGGAA
AGCGAAAGAA GGGAAGGAAA

2451 CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC GGGGGCTCCC
TTTAGGGTTC CGATTTAGTG
          GAGCGGTGCA AGCGGCCGAA AGGGGCAGTT CGAGATTTAG CCCCCGAGGG
AAATCCCAAG GCTAAATCAC

2521 CTTTACGGCA CCTCGACCCC AAAAAACTTG ATTAGGGTGA TGGTTCACGT
AGTGGGCCAT CGCCCTGATA
          GAAATGCCGT GGAGCTGGGG TTTTTTGAAC TAATCCCACT ACCAAGTGCA
TCACCCGGTA GCGGGACTAT

2591 GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT AATAGTGGAC
TCTTGTTCCA AACTGGAACA
          CTGCCAAAAA GCGGGAAACT GCAACCTCAG GTGCAAGAAA TTATCACCTG
AGAACAAGGT TTGACCTTGT

2661 ACACTCAACC CTATCTCGGT CTATTCTTTT GATTTATAAG GGATTTTGCC
GATTTCGGCC TATTGGTTAA
          TGTGAGTTGG GATAGAGCCA GATAAGAAAA CTAAATATTC CCTAAAACGG
CTAAAGCCGG ATAACCAATT

2731 AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATTA
ACGCTTACAA TTTAGGTGGC
          TTTTACTCGA CTAAATTGTT TTTAAATTGC GCTTAAAATT GTTTTATAAT
TGCGAATGTT AAATCCACCG

2801 ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT
ACATTCAAAT ATGTATCCGC
          TGAAAGCCC CTTTACACGC GCCTTGGGGA TAAACAAATA AAAAGATTTA
TGTAAGTTTA TACATAGGCG

2871 TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG AAAAAGGAAG
AGTATGAGTA TTCAACATTT
          AGTACTCTGT TATTGGGACT ATTTACGAAG TTATTATAAC TTTTTCCTTC
TCATACTCAT AAGTTGTAAA

2941 CCGTGTCGCC CTTATTCCCT TTTTTGCGGC ATTTTGCCTT CCTGTTTTTG
CTCACCCAGA AACGCTGGTG
          GGCACAGCGG GAATAAGGGA AAAAACGCCG TAAAACGGAA GGACAAAAAC
GAGTGGGTCT TTGCGACCAC

3011 AAAGTAAAAG ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA
ACTGGATCTC AACAGCGGTA
          TTTCATTTTC TACGACTTCT AGTCAACCCA CGTGCTCACC CAATGTAGCT
TGACCTAGAG TTGTCGCCAT

Fig. 48 F

3081   AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT
TTTAAAGTTC TGCTATGTGG
        TCTAGGAACT CTCAAAAGCG GGGCTTCTTG CAAAAGGTTA CTACTCGTGA
AAATTTCAAG ACGATACACC

3151   CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC GGTCGCCGCA
TACACTATTC TCAGAATGAC
        GCGCCATAAT AGGGCATAAC TGCGGCCCGT TCTCGTTGAG CCAGCGGCGT
ATGTGATAAG AGTCTTACTG

3221   TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC
AGTAAGAGAA TTATGCAGTG
        AACCAACTCA TGAGTGGTCA GTGTCTTTTC GTAGAATGCC TACCGTACTG
TCATTCTCTT AATACGTCAC

3291   CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACAACG
ATCGGAGGAC CGAAGGAGCT
        GACGGTATTG GTACTCACTA TTGTGACGCC GGTTGAATGA AGACTGTTGC
TAGCCTCCTG GCTTCCTCGA

3361   AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC CTTGATCGTT
GGGAACCGGA GCTGAATGAA
        TTGGCGAAAA AACGTGTTGT ACCCCCTAGT ACATTGAGCG GAACTAGCAA
CCCTTGGCCT CGACTTACTT

3431   GCCATACCAA ACGACGAGCG TGACACCACG ATGCCTGTAG CAATGGCAAC
AACGTTGCGC AAACTATTAA
        CGGTATGGTT TGCTGCTCGC ACTGTGGTGC TACGGACATC GTTACCGTTG
TTGCAACGCG TTTGATAATT

3501   CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAATTAAT AGACTGGATG
GAGGCGGATA AAGTTGCAGG
        GACCGCTTGA TGAATGAGAT CGAAGGGCCG TTGTTAATTA TCTGACCTAC
CTCCGCCTAT TTCAACGTCC

3571   ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT
CTGGAGCCGG TGAGCGTGGG
        TGGTGAAGAC GCGAGCCGGG AAGGCCGACC GACCAAATAA CGACTATTTA
GACCTCGGCC ACTCGCACCC

3641   TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT
CGTAGTTATC TACACGACGG
        AGAGCGCCAT AGTAACGTCG TGACCCCGGT CTACCATTCG GGAGGGCATA
GCATCAATAG ATGTGCTGCC

3711   GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT
GCCTCACTGA TTAAGCATTG
        CCTCAGTCCG TTGATACCTA CTTGCTTTAT CTGTCTAGCG ACTCTATCCA
CGGAGTGACT AATTCGTAAC

3781   GTAACTGTCA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAAC
TTCATTTTTA ATTTAAAAGG

Fig. 48 G

```
            CATTGACAGT CTGGTTCAAA TGAGTATATA TGAAATCTAA CTAAATTTTG
AAGTAAAAAT TAAATTTTCC

    3851  ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA TCCCTTAACG
TGAGTTTTCG TTCCACTGAG
            TAGATCCACT TCTAGGAAAA ACTATTAGAG TACTGGTTTT AGGGAATTGC
ACTCAAAAGC AAGGTGACTC

    3921  CGTCAGACCC CGTAGAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT
CTGCGCGTAA TCTGCTGCTT
            GCAGTCTGGG GCATCTTTTC TAGTTTCCTA GAAGAACTCT AGGAAAAAAA
GACGCGCATT AGACGACGAA

    3991  GCAAACAAAA AAACCACCGC TACCAGCGGT GGTTTGTTTG CCGGATCAAG
AGCTACCAAC TCTTTTTCCG
            CGTTTGTTTT TTTGGTGGCG ATGGTCGCCA CCAAACAAAC GGCCTAGTTC
TCGATGGTTG AGAAAAAGGC

    4061  AAGGTAACTG GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT
GTAGCCGTAG TTAGGCCACC
            TTCCATTGAC CGAAGTCGTC TCGCGTCTAT GGTTTATGAC AGGAAGATCA
CATCGGCATC AATCCGGTGG

    4131  ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG
TTACCAGTGG CTGCTGCCAG
            TGAAGTTCTT GAGACATCGT GGCGGATGTA TGGAGCGAGA CGATTAGGAC
AATGGTCACC GACGACGGTC

    4201  TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG
ATAAGGCGCA GCGGTCGGGC
            ACCGCTATTC AGCACAGAAT GGCCCAACCT GAGTTCTGCT ATCAATGGCC
TATTCCGCGT CGCCAGCCCG

    4271  TGAACGGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA CGACCTACAC
CGAACTGAGA TACCTACAGC
            ACTTGCCCCC CAAGCACGTG TGTCGGGTCG AACCTCGCTT GCTGGATGTG
GCTTGACTCT ATGGATGTCG

    4341  GTGAGCTATG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GGCGGACAGG
TATCCGGTAA GCGGCAGGGT
            CACTCGATAC TCTTTCGCGG TGCGAAGGGC TTCCCTCTTT CCGCCTGTCC
ATAGGCCATT CGCCGTCCCA

    4411  CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAAC GCCTGGTATC
TTTATAGTCC TGTCGGGTTT
            GCCTTGTCCT CTCGCGTGCT CCCTCGAAGG TCCCCCTTTG CGGACCATAG
AAATATCAGG ACAGCCCAAA

    4481  CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG
GAGCCTATGG AAAAACGCCA
            GCGGTGGAGA CTGAACTCGC AGCTAAAAAC ACTACGAGCA GTCCCCCCGC
CTCGGATACC TTTTTGCGGT
```

Fig. 48 H

4551 GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC TTTTGCTCAC
ATGTTCTTTC CTGCGTTATC
        CGTTGCGCCG GAAAAATGCC AAGGACCGGA AAACGACCGG AAAACGAGTG
TACAAGAAAG GACGCAATAG

4621 CCCTGATTCT GTGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACCG
CTCGCCGCAG CCGAACGACC
        GGGACTAAGA CACCTATTGG CATAATGGCG GAAACTCACT CGACTATGGC
GAGCGGCGTC GGCTTGCTGG

4691 GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA
ACCGCCTCTC CCCGCGCGTT
        CTCGCGTCGC TCAGTCACTC GCTCCTTCGC CTTCTCGCGG GTTATGCGTT
TGGCGGAGAG GGGCGCGCAA

4761 GGCCGATTCA TTAATGCAGC TGGCACGACA GGTTTCCCGA CTGGAAAGCG
GGCAGTGAGC GCAACGCAAT
        CCGGCTAAGT AATTACGTCG ACCGTGCTGT CCAAAGGGCT GACCTTTCGC
CCGTCACTCG CGTTGCGTTA

4831 TAATGTGAGT TAGCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATGC
TTCCGGCTCG TATGTTGTGT
        ATTACACTCA ATCGAGTGAG TAATCCGTGG GGTCCGAAAT GTGAAATACG
AAGGCCGAGC ATACAACACA

BssHII

~~~~~~~
4901 GGAATTGTGA GCGGATAACA ATTTCACACA GGAAACAGCT ATGACCATGA
TTACGCCAAG CGCGCAATTA
        CCTTAACACT CGCCTATTGT TAAAGTGTGT CCTTTGTCGA TACTGGTACT
AATGCGGTTC GCGCGTTAAT

                                        NcoI

                                        KpnI
                                        ~~~~~
4971 ACCCTCACTA AAGGGAACAA AAGCTGGGTA C
        TGGGAGTGAT TTCCCTTGTT TTCGACCCAT G

Fig. 48 I

Figure 49 A

Rice actin

Bgl II (5271)

Bam HI (5239)

Hin dIII (2)

Nco I (9)

aleurain signal/NPIR

NPIR

Not I (129)

Sal I (229)

Bam HI (285)   FAE

Xho I (667)

Nco I (754)

Eco RI (4537)

Xho I (4161)

Kpn I (4150)

pTP11-1

5387 bp

Native COOH

nos

ori

Cla I (1261)

Hin dIII (1266)

Xba I (1271)

f1 ori

amp

Figure 49B

```
              NcoI
              ~~~~~~
      HindIII
      ~~~~~~
              M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A   V
   1    AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG
                                                                          NotI
                                                                          ~~~~~~~
         ·  A   S   S   S   S   F   A   D   S   N   P   I   R   P   V   T   D   R   A   A   A   S   T ·
  71    TCGCCTCCTC CTCCTCCTTC GCCGACTCCA ACCCGATCCG GCCCGTCACC GACCGCGCGG CCGCCTCCAC
         · Q   G   I   S   E   D   L   Y   S   R   L   V   E   M   A   T   I   S   Q   A   A   Y   A
 141    GCAGGGCATC TCCGAAGACC TCTACAGCCG TTTAGTCGAA ATGGCCACTA TCTCCCAAGC TGCCTACGCC
                    SalI
                    ~~~~~~~
                    AccI
                    ~~~~~~~
              D   L   C   N   I   P   S   T   I   I   K   G   E   K   I   Y   N   S   Q   T   D   I   N   G
 211    GACCTGTGCA ACATTCCGTC GACTATTATC AAGGGAGAGA AAATTTACAA TTCTCAAACT GACATTAACG
              BamHI
              ~~~~~~
         ·  W   I   L   R   D   D   S   S   K   E   I   I   T   V   F   R   G   T   G   S   D   T   N ·
 281    GATGGATCCT CCGCGACGAC AGCAGCAAAG AAATAATCAC CGTCTTCCGT GGCACTGGTA GTGATACGAA
         · L   Q   L   D   T   N   Y   T   L   T   P   F   D   T   L   P   Q   C   N   G   C   E   V
 351    TCTACAACTC GATACTAACT ACACCCTCAC GCCTTTCGAC ACCCTACCAC AATGCAACGG TTGTGAAGTA
          H   G   G   Y   Y   I   G   W   V   S   V   Q   D   Q   V   E   S   L   V   K   Q   Q   V   S
 421    CACGGTGGAT ATTATATTGG ATGGGTCTCC GTCCAGGACC AAGTCGAGTC GCTTGTCAAA CAGCAGGTTA
         · Q   Y   P   D   Y   A   L   T   V   T   G   H   X   L   G   A   S   L   A   A   L   T   A ·
 491    GCCAGTATCC GGACTACGCG CTGACCGTGA CCGGCCACKC CCTCGGCGCC TCCCTGGCGG CACTCACTGC
         · A   Q   L   S   A   T   Y   D   N   I   R   L   Y   T   F   G   E   P   R   S   G   N   Q
 561    CGCCCAGCTG TCTGCGACAT ACGACAACAT CCGCCTGTAC ACCTTCGGCG AACCGCGCAG CGGCAATCAG
                                                            XhoI
                                                            ~~~~~~~
              A   F   A   S   Y   M   N   D   A   F   Q   A   S   S   P   D   T   T   Q   Y   F   R   V   T
 631    GCCTTCGCGT CGTACATGAA CGATGCCTTC CAAGCCTCGA GCCCAGATAC GACGCAGTAT TTCCGGGTCA
                                                              NcoI
                                                              ~~~~~~~
         ·  H   A   N   D   G   I   P   N   L   P   P   V   E   Q   G   Y   A   H   G   G   V   E   Y ·
 701    CTCATGCCAA CGACGGCATC CCAAACCTGC CCCCGGTGGA GCAGGGGTAC GCCCATGGCG GTGTAGAGTA
         · W   S   V   D   P   Y   S   A   Q   N   T   F   V   C   T   G   D   E   V   Q   C   C   E
 771    CTGGAGCGTT GATCCTTACA GCGCCCAGAA CACATTTGTC TGCACTGGGG ATGAAGTGCA GTGCTGTGAG
          A   Q   G   G   Q   G   V   N   N   A   H   T   T   Y   F   G   M   T   S   G   A   C   T   W
 841    GCCCAGGGCG GACAGGGTGT GAATAATGCG CACACGACTT ATTTTGGGAT GACGAGCGGA GCCTGTACAT
         ·   *
 911    GGTGATCAGT CATTTCAGCC TCCCCGAGTG TACCAGGAAA GATGGATGTC CTGGAGAGGG GGCCGCGTAA
 981    CCACTGAAGG ATGAGCTGTA AAGAAGCAGA TCGTTCAAAC ATTTGGCAAT AAAGTTTCTT AAGATTGAAT
1051    CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT TGAATTACGT TAAGCATGTA ATAATTAACA
1121    TGTAATGCAT GACGTTATTT ATGAGATGGG TTTTTATGAT TAGAGTCCCG CAATTATACA TTTAATACGC
                                                                                   ClaI
1191    GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA TCGCGCGCGG TGTCATCTAT GTTACTAGAT
              HindIII
              ~~~~~~~
        ClaI       XbaI
        ~~~~       ~~~~~~~
1261    CGATAAGCTT CTAGAGCGGC CGGTGGAGCT CCAATTCGCC CTATAGTGAG TCGTATTACG CGCGCTCACT
1331    GGCCGTCGTT TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT
1401    CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG TTGCGCAGCC
1471    TGAATGGCGA ATGGGACGCG CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT
1541    GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT CGCCACGTTC
1611    GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC
1681    TCGACCCCAA AAAACTTGAT TAGGGTGATG GTTCACGTAG TGGGCCATCG CCCTGATAGA CGGTTTTTCG
```
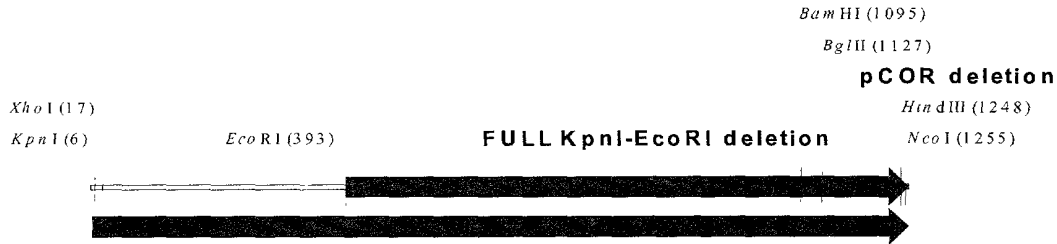
Fig. 49 C

```
1751    CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCT
1821    ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA
1891    TTTAACAAAA ATTTAACGCG AATTTTAACA AAATATTAAC GCTTACAATT TAGGTGGCAC TTTTCGGGGA
1961    AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
2031    AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT
2101    TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT
2171    GCTGAAGATC AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA
2241    GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC
2311    CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC
2381    TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA
2451    TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT
2521    GCACAACATG GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC
2591    GACGAGCGTG ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC
2661    TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC CACTTCTGCG
2731    CTCGGCCCTT CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC
2801    ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA
2871    CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA
2941    CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT CATTTTTAAT TTAAAAGGAT CTAGGTGAAG
3011    ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG
3081    TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA
3151    ACCACCGCTA CCAGCGGTGG TTTGTTTGCC GGATCAAGAG CTACCAACTC TTTTTCCGAA GGTAACTGGC
3221    TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT
3291    CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC
3361    GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT
3431    TCGTGCACAC AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA CCTACAGCGT GAGCTATGAG
3501    AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA
3571    GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA
3641    CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCCT
3711    TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT
3781    GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGCAGCC GAACGACCGA GCGCAGCGAG
3851    TCAGTGAGCG AGGAAGCGGA AGAGCGCCCA ATACGCAAAC CGCCTCTCCC CGCGCGTTGG CCGATTCATT
3921    AATGCAGCTG GCACGACAGG TTTCCCGACT GGAAAGCGGG CAGTGAGCGC AACGCAATTA ATGTGAGTTA
3991    GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT CCGGCTCGTA TGTTGTGTGG AATTGTGAGC
4061    GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT ACGCCAAGCG CGCAATTAAC CCTCACTAAA
                            KpnI                   XhoI
                            ~~~~~~~                ~~~~~~
4131    GGGAACAAAA GCTGGGTACC GGGCCCCCCC TCGAGGTCAT TCATATGCTT GAGAAGAGAG TCGGGATAGT
4201    CCAAAATAAA ACAAAGGTAA GATTACCTGG TCAAAAGTGA AAACATCAGT TAAAAGGTGG TATAAGTAAA
4271    ATATCGGTAA TAAAAGGTGG CCCAAAGTGA AATTTACTCT TTTCTACTAT TATAAAAATT GAGGATGTTT
4341    TGTCGGTACT TTGATACGTC ATTTTTGTAT GAATTGGTTT TTAAGTTTAT TCGCGATTTG GAAATGCATA
4411    TCTGTATTTG AGTCGGTTTT TAAGTTCGTT GCTTTTGTAA ATACAGAGGG ATTTGTATAA GAAATATCTT
                                                                            EcoRI
                                                                            ~~~~~~
4481    TAAAAAACCC ATATGCTAAT TTGACATAAT TTTTGAGAAA AATATATATT CAGGCGAATT CCACAATGAA
4551    CAATAATAAG ATTAAAATAG CTTGCCCCCG TTGCAGCGAT GGGTATTTTT TCTAGTAAAA TAAAAGATAA
4621    ACTTAGACTC AAAACATTTA CAAAACAAC CCCTAAAGTC CTAAAGCCCA AAGTGCTATG CACGATCCAT
4691    AGCAAGCCCA GCCCAACCCA ACCCAACCCA ACCCACCCCA GTGCAGCCAA CTGGCAAATA GTCTCCACCC
4761    CCGGCACTAT CACCGTGAGT TGTCCGCACC ACCGCACGTC TCGCAGCCAA AAAAAAAAA AGAAAGAAAA
4831    AAAAGAAAAA GAAAAACAGC AGGTGGGTCC GGGTCGTGGG GGCCGGAAAA GCGAGGAGGA TCGCGAGCAG
4901    CGACGAGGCC CGGCCCTCCC TCCGCTTCCA AAGAAACGCC CCCATCGCC ACTATATACA TACCCCCCCC
4971    TCTCCTCCCA TCCCCCCAAC CCTACCACCA CCACCACCAC CACCTCCTCC CCCCTCGCTG CCGGACGACG
5041    AGCTCCTCCC CCCTCCCCCT CCGCCGCCGC CGGTAACCAC CCCGCCCCTC TCCTCTTTCT TTCTCCGTTT
5111    TTTTTTTCGT CTCGGTCTCG ATCTTTGGCC TTGGTAGTTT GGGTGGGCGA GAGCGGCTTC GTCGCCCAGA
                                                                            BamHI
                                                                            ~~~~~~
5181    TCGGTGCGCG GGAGGGGCGG GATCTCGCGG CTGGCGTCTC CGGGCGTGAG TCGGCCCGGA TCCTCGCGGG
                            BglII
                            ~~~~~~
5251    GAATGGGGCT CTCGGATGTA GATCTTCTTT CTTTCTTCTT TTTGTGGTAG AATTTGAATC CCTCAGCATT
5321    GTTCATCGGT AGTTTTTCTT TTCATGATTT GTGACAAATG CAGCCTCGTG CGGAGCTTTT TTGTAGC
```

Figure 50 A

Actin promoter -FAEs

Actin promoter -FAEs map showing restriction sites: Xho I (17), Kpn I (6), Eco R1 (393), Bam HI (1095), Bgl II (1127), Hin d III (1248), Nco I (1255). FULL Kpnl-EcoRl deletion, pCOR deletion.

ACTIN-PROMOTER-FAEs
1259 bp

```
            KpnI                    XhoI
            ~~~~~~                   ~~~~~~
      1  GGTACCGGGC CCCCCCTCGA GGTCATTCAT ATGCTTGAGA AGAGAGTCGG GATAGTCCAA AATAAAACAA
         CCATGGCCCG GGGGGGAGCT CCAGTAAGTA TACGAACTCT TCTCTCAGCC CTATCAGGTT TTATTTTGTT

     71  AGGTAAGATT ACCTGGTCAA AAGTGAAAAC ATCAGTTAAA AGGTGGTATA AGTAAAATAT CGGTAATAAA
         TCCATTCTAA TGGACCAGTT TTCACTTTTG TAGTCAATTT TCCACCATAT TCATTTTATA GCCATTATTT

    141  AGGTGGCCCA AAGTGAAATT TACTCTTTTC TACTATTATA AAAATTGAGG ATGTTTTGTC GGTACTTTGA
         TCCACCGGGT TTCACTTTAA ATGAGAAAAG ATGATAATAT TTTTAACTCC TACAAAACAG CCATGAAACT

    211  TACGTCATTT TTGTATGAAT TGGTTTTTAA GTTTATTCGC GATTTGGAAA TGCATATCTG TATTTGAGTC
         ATGCAGTAAA AACATACTTA ACCAAAAATT CAAATAAGCG CTAAACCTTT ACGTATAGAC ATAAACTCAG

    281  GGTTTTTAAG TTCGTTGCTT TTGTAAATAC AGAGGGATTT GTATAAGAAA TATCTTTAAA AAACCCATAT
         CCAAAAATTC AAGCAACGAA AACATTTATG TCTCCCTAAA CATATTCTTT ATAGAAATTT TTTGGGTATA

                                                   EcoRI
                                                   ~~~~~~
    351  GCTAATTTGA CATAATTTTT GAGAAAAATA TATATTCAGG CGAATTCCAC AATGAACAAT AATAAGATTA
         CGATTAAACT GTATTAAAAA CTCTTTTTAT ATATAAGTCC GCTTAAGGTG TTACTTGTTA TTATTCTAAT

    421  AAATAGCTTG CCCCCGTTGC AGCGATGGGT ATTTTTTCTA GTAAAATAAA AGATAAACTT AGACTCAAAA
         TTTATCGAAC GGGGGCAACG TCGCTACCCA TAAAAAAGAT CATTTTATTT TCTATTTGAA TCTGAGTTTT

    491  CATTTACAAA AACAACCCCT AAAGTCCTAA AGCCCAAAGT GCTATGCACG ATCCATAGCA AGCCCAGCCC
         GTAAATGTTT TTGTTGGGGA TTTCAGGATT TCGGGTTTCA CGATACGTGC TAGGTATCGT TCGGGTCGGG

    561  AACCCAACCC AACCCAACCC ACCCCAGTGC AGCCAACTGG CAAATAGTCT CCACCCCCGG CACTATCACC
         TTGGGTTGGG TTGGGTTGGG TGGGGTCACG TCGGTTGACC GTTTATCAGA GGTGGGGGCC GTGATAGTGG

    631  GTGAGTTGTC CGCACCACCG CACGTCTCGC AGCCAAAAAA AAAAAAAGAA AGAAAAAAAA GAAAAAGAAA
         CACTCAACAG GCGTGGTGGC GTGCAGAGCG TCGGTTTTTT TTTTTTTCTT TCTTTTTTTT CTTTTTCTTT

    701  AACAGCAGGT GGGTCCGGGT CGTGGGGGCC GGAAAAGCGA GGAGGATCGC GAGCAGCGAC GAGGCCCGGC
         TTGTCGTCCA CCCAGGCCCA GCACCCCCGG CCTTTTCGCT CCTCCTAGCG CTCGTCGCTG CTCCGGGCCG
```

Figure 50 B

```
771   CCTCCCTCCG CTTCCAAAGA AACGCCCCCC ATCGCCACTA TATACATACC CCCCCCTCTC CTCCCATCCC
      GGAGGGAGGC GAAGGTTTCT TTGCGGGGGG TAGCGGTGAT ATATGTATGG GGGGGGAGAG GAGGGTAGGG

841   CCCAACCCTA CCACCACCAC CACCACCACC TCCTCCCCCC TCGCTGCCGG ACGACGAGCT CCTCCCCCCT
      GGGTTGGGAT GGTGGTGGTG GTGGTGGTGG AGGAGGGGGG AGCGACGGCC TGCTGCTCGA GGAGGGGGGA

911   CCCCCTCCGC CGCCGCCGGT AACCACCCCG CCCCTCTCCT CTTTCTTTCT CCGTTTTTTT TTTCGTCTCG
      GGGGGAGGCG GCGGCGGCCA TTGGTGGGGC GGGGAGAGGA GAAAGAAAGA GGCAAAAAAA AAAGCAGAGC

981   GTCTCGATCT TTGGCCTTGG TAGTTTGGGT GGGCGAGAGC GGCTTCGTCG CCCAGATCGG TGCGCGGGAG
      CAGAGCTAGA AACCGGAACC ATCAAACCCA CCCGCTCTCG CCGAAGCAGC GGGTCTAGCC ACGCGCCCTC

                                                  BamHI
                                                  ~~~~~
1051  GGGCGGGATC TCGCGGCTGG CGTCTCCGGG CGTGAGTCGG CCCGGATCCT CGCGGGGAAT GGGGCTCTCG
      CCCGCCCTAG AGCGCCGACC GCAGAGGCCC GCACTCAGCC GGGCCTAGGA GCGCCCCTTA CCCCGAGAGC

          BglII
          ~~~~~~~
1121  GATGTAGATC TTCTTTCTTT CTTCTTTTTG TGGTAGAATT TGAATCCCTC AGCATTGTTC ATCGGTAGTT
      CTACATCTAG AAGAAAGAAA GAAGAAAAAC ACCATCTTAA ACTTAGGGAG TCGTAACAAG TAGCCATCAA

                                                             HindIII  NcoI
                                                             ~~~~~~~  ~~~~~
1191  TTTCTTTTCA TGATTTGTGA CAAATGCAGC CTCGTGCGGA GCTTTTTTGT AGGTAGAAGC TTACCATGG
      AAAGAAAAGT ACTAAACACT GTTTACGTCG GAGCACGCCT CGAAAAAACA TCCATCTTCG AATGGTACC
```

Kpn1-EcoR1 - deletion underlined and restored NCO site in bold in vectors pJQ4.9, pJQ3.2 and pJO6.3.

Figure 5/

ALEURAIN_deleted NPIR (Apoplast) structure and sequence

Ncol(9)

HindIII(2)   ALE-CUT   NotI(87)

ALEURAIN-NPIR-DEL
93 bp

```
+1              M   A   H   A   R   V   L   L   L   A   L   A   V   L   A   T   A   A   V   A
        HindIII NcoI
        ~~~~~~ ~~~~~~~
1    AAGCTTACCA TGGCCCACGC CCGCGTCCTC CTCCTGGCGC TCGCCGTGCT GGCCACGGCC GCCGTCGCCG
     TTCGAATGGT ACCGGGTGCG GGCGCAGGAG GAGGACCGCG AGCGGCACGA CCGGTGCCGG CGGCAGCGGC


+1 V   A   S   S   R   A   A
                     NotI
                     ~~~~~~~~~
71   TCGCCTCCTC CCGCGCGGCC GCC
     AGCGGAGGAG GGCGCGCCGG CGG
```

Figure 52

SEE1 (Senescence enhanced) PROMOTER sequence

```
  1    CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
                         PstI                                 XbaI      XbaI
                        ~~~~~~~                              ~~~~~~~ ~~~~~~
 71    AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTTGT ATCTAGAACT CTAGATAGCA CAGCCACAGC
141    ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTTG GAGACGGAGC TCTTTCCTAC CTCCTGACGT
211    TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC CCAACAAAAT ATCGTCCCCC
281    ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTCGCGCC GTTTTTGTCT GAATCTCGCT TCCACTGGCC
                     SmaI
                    ~~~~~~
351    AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCG TCACCCCTGG CGTCATGGGA
421    TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCAAT
                                        SphI
                                       ~~~~~~~
491    TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTTTGG CTTAATTGAC TTTATTTTTG
561    TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
631    CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTACA GCTGGAGGCT
                          PstI
                         ~~~~~~~
701    TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
771    CCCCGGCACC GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCTCGC TCCCCCTGCC GGACGACCCA
841    GTAAAATACT GTTGCCCACT CGCCGGCGAG ATG
```

Figure 53

SEE1 (Senescence enhanced) PROMOTER plus vacuolar aleurain SIGNAL/NPIR sequence

```
  1    CATGGGCCAG GTATAATTAT GGGATATCTC AAGCAAATAA TCGAAATATC ACCATTGGCT ACAATATCTG
                     PstI                              XbaI      XbaI
                     ~~~~~~~                           ~~~~~~~   ~~~~~~
 71    AGCTCCGAGT TCTGACTGCA GTCTGGATGA CGCGTGTTGT ATCTAGAACT CTAGATAGCA CAGCCACAGC
141    ACCTACAGGA GTGCGACACT TGTGGACTGT AGTAGTGTTG GAGACGGAGC TCTTTCCTAC CTCCTGACGT
211    TGCCGCCGTT GTCCATTCCA ACGGCATCAC TCTCAACCAA TCACGCGCTC CCAACAAAAT ATCGTCCCCC
281    ATGTCTTGGC GGAGAGAGAG TACATACATG CTGTCGCGCC GTTTTTGTCT GAATCTCGCT TCCACTGGCC
                     SmaI
                     ~~~~~~~
351    AATCAGCTCA GCTCCCGGGA GCTCACTCAT TCAAGATCCC ATCGTCGTCG TCACCCCTGG CGTCATGGGA
421    TGGAAAAGAA CCTCCGTTGC TCGGATGAGT CAGCCATATC CCCGAACAGA GTACTGCAAG ATAACCCAAT
                               SphI
                               ~~~~~~~
491    TCAGATTCCC CCAATAGAGA AAGTATAGCA TGCTTTCGGG TTTTGTTTGG CTTAATTGAC TTTATTTTTG
561    TTGGAGTTGA ATGCTGATTT GTTGTGTAAA ATGCCCAACC ATCTGAATAT CGAGACGGAT AATAGGCTGG
631    CTAATTAATT TATAGCAAGA TTCTGTAGTG CACATCGCAA ATATCTTTCT GGGCATTACA GCTGGAGGCT
                     PstI
                     ~~~~~~~
701    TCATCAGCCT GAAACACTCT GCAGAGCCTG AAGCAAGTGG TGAAGCGTGG CGATGAGATG GGTATAAAAC
771    CCCCGGCACC GGGACGCGAG CTCCCGCCTA CCAGTACCAT CTCGCCTCGC TCCCCCTGCC GGACGACCCA
                                            M  A  H  G   R  I  L   F  L  A   L  A  V  L
841    GTAAAATACT GTTGCCCACT CGCCGGCGAG ATGGCCCACG CCGCATCCT CTTCTTGGCG CTCGCCGTCT
                                                            BssHII
                                                                    NotI
       · A  T  A   A  V  A   A  A  S   L  A  D   S  N  P   I  R  P   V  T   E  R  A ·
911    TGGCCACCGC CGCGGTGGCC GCCGCATCNT TGGCGGACTC CAACCCGATC CGGCCCGTCA CCGAGCGCGC
       NotI
       ~~~~~~
       · A  A
981    GGCCGCC
```